Spatial Analysis of Maritime Traffic for Maritime Security

Ronald Pelot, Ph.D., P.Eng.

Research assistants

Dong Lin, Ph.D. Yan Wu, Ph.D.

Casey Hilliard David Wootton

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **OCT 2009**	2. REPORT TYPE	3. DATES COVERED **00-00-2009 to 00-00-2009**

4. TITLE AND SUBTITLE **Spatial Analysis of Maritime Traffic for Maritime Security**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Dalhousie University,Networks of Centres of Excellence (NSERC),Halifax, Nova Scotia, Canada B3H 4R2,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Maritime Domain Awareness and Counter Piracy, 26-29 October 2009, Ottawa, Canada. U.S. Government or Federal Rights License

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**58**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Traffic Modelling

Activity Types

Commercial shipping
Commercial fishing
Commercial recreational
Aquaculture
Ferries
Cruise ships
Private recreational

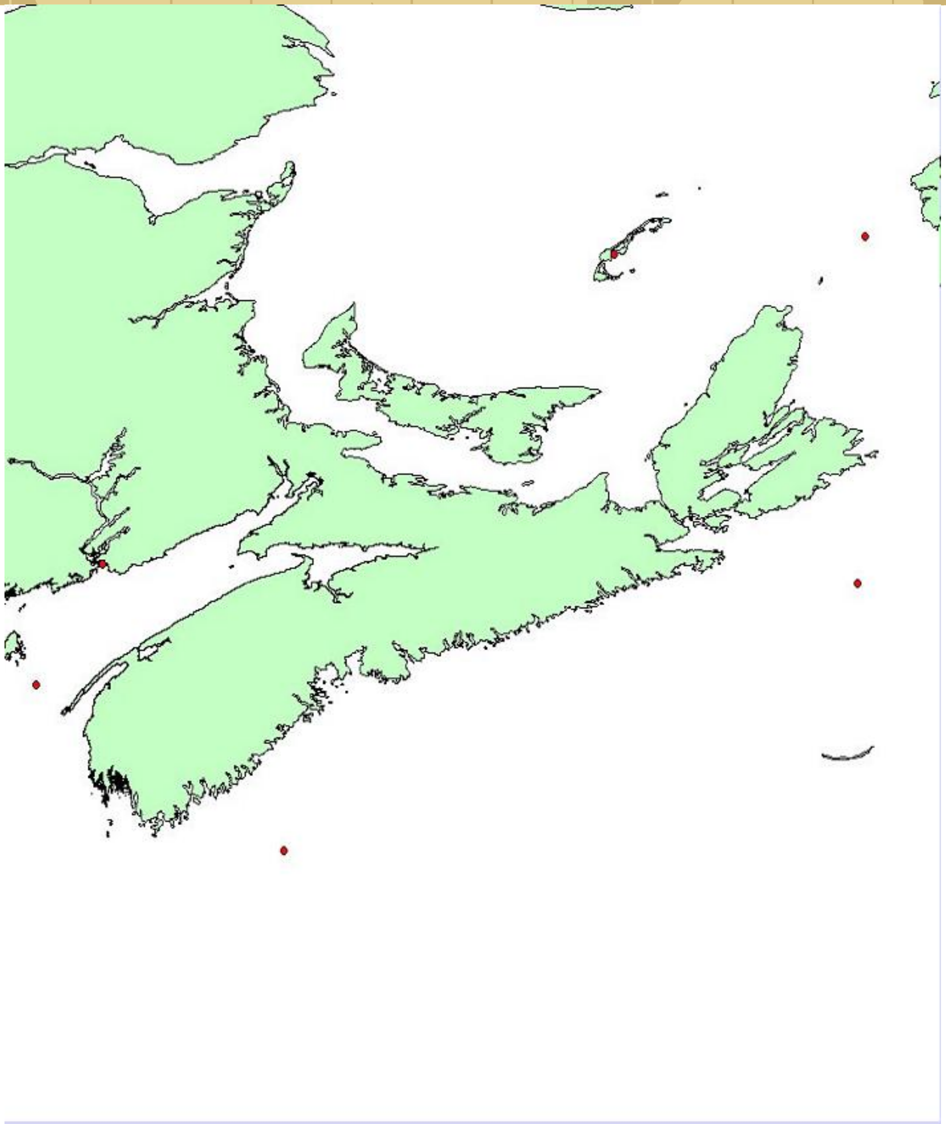
Data quality

Class	Available Information	Example
A	Continuous time-stamped position data; vessel info	Canadian West Coast shipping transits; Automatic Identification System (AIS)
B	Specific origin & destination, and intermediate waypoints; travel dates; vessel info	Canadian East Coast shipping transits
C	Specific origin & destination; travel dates; vessel info	Ferries; Cruise ships
D	Specific origin; general destination; travel dates; vessel info	Commercial fishing by NAFO zones
E	Specific origin; general destination; frequency of trips; vessel info	Lobster fishing; ecotours
F	General origin; general destination; frequency of trips; classes of vessels	Recreational boating

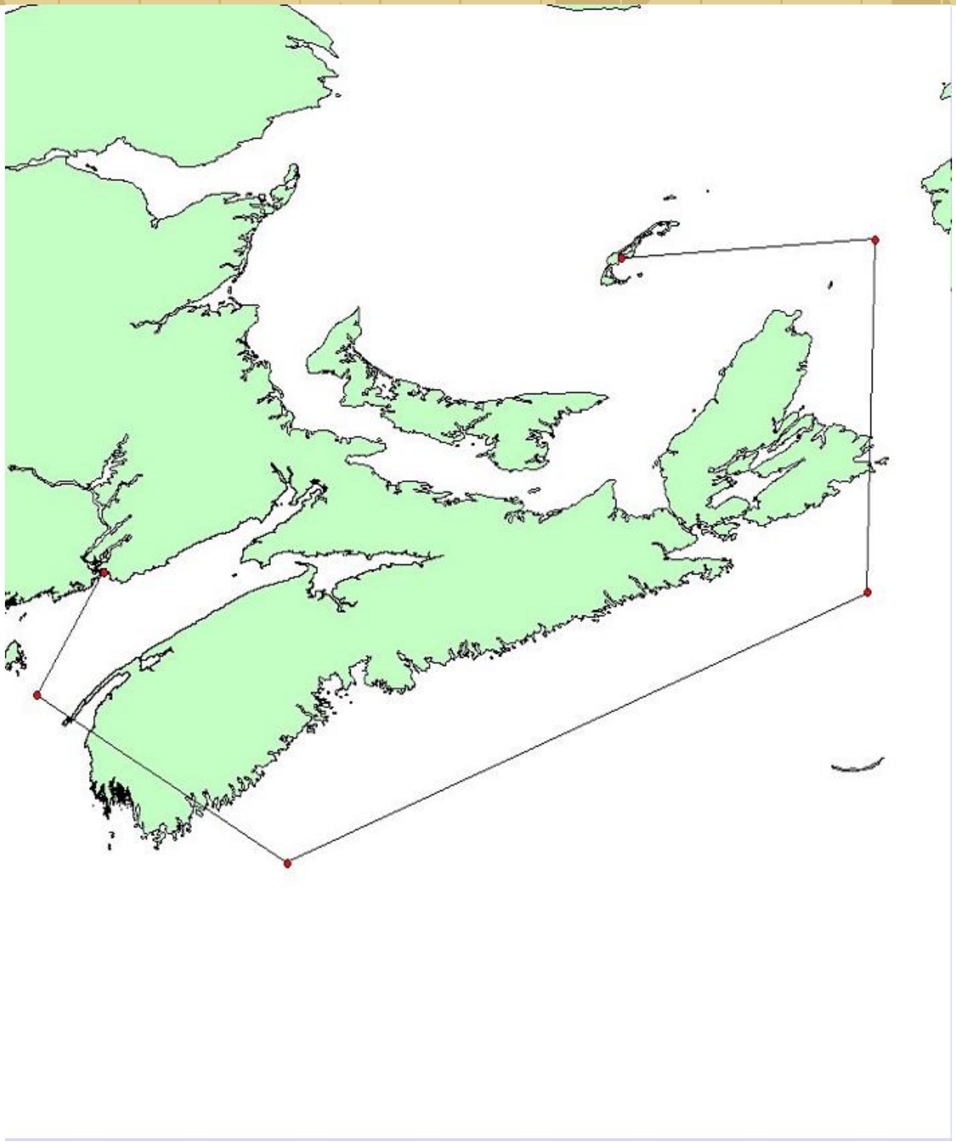
Traffic Modelling

- Maximize use of information:
  - source: location, date, time
  - intermediate points: way points, fishing efforts
  - final destination: location, date, time
- Extrapolate using more general information
  - feasible range: distance, time
  - feasible location; eg. fishing grounds; tourist sites
  - typical locations: cluster analysis

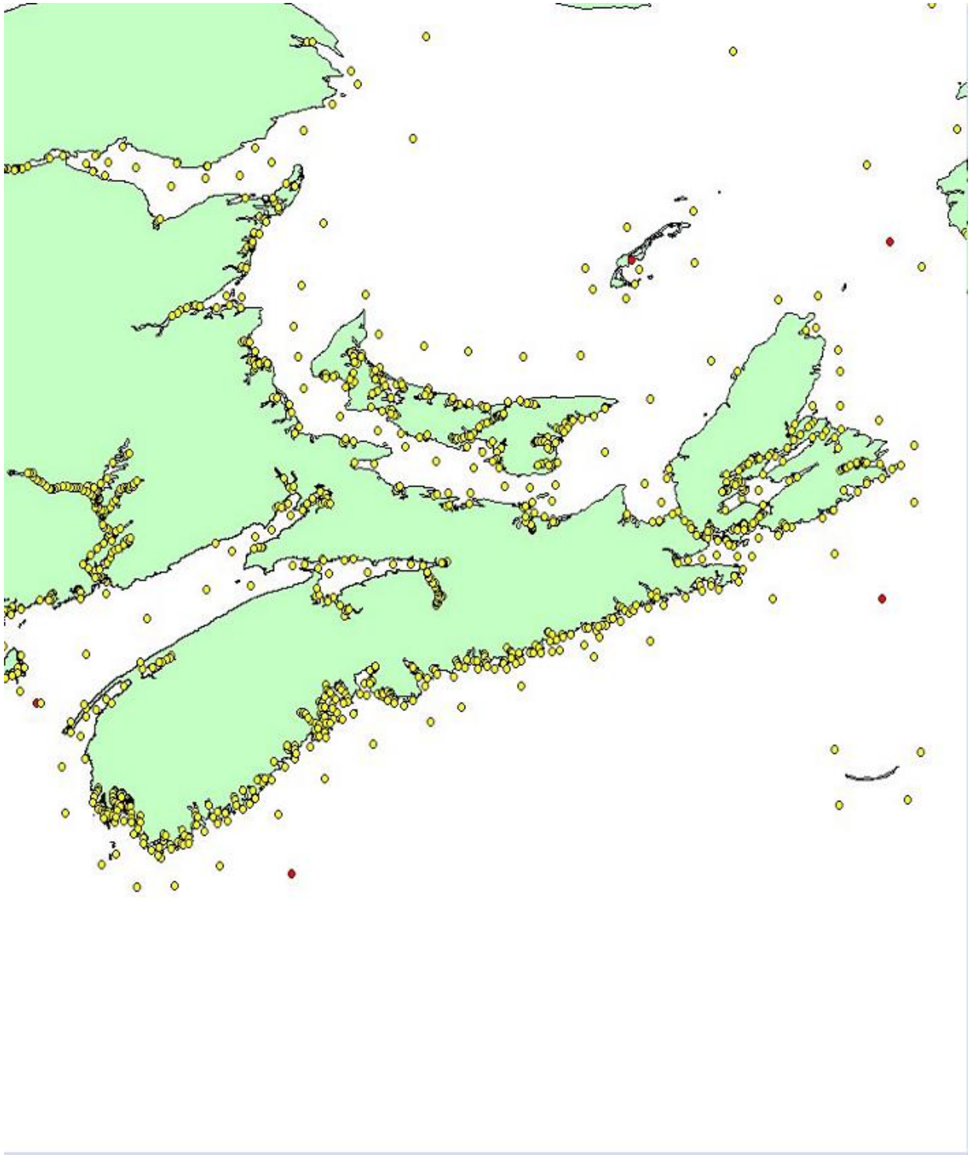
Trip Waypoints



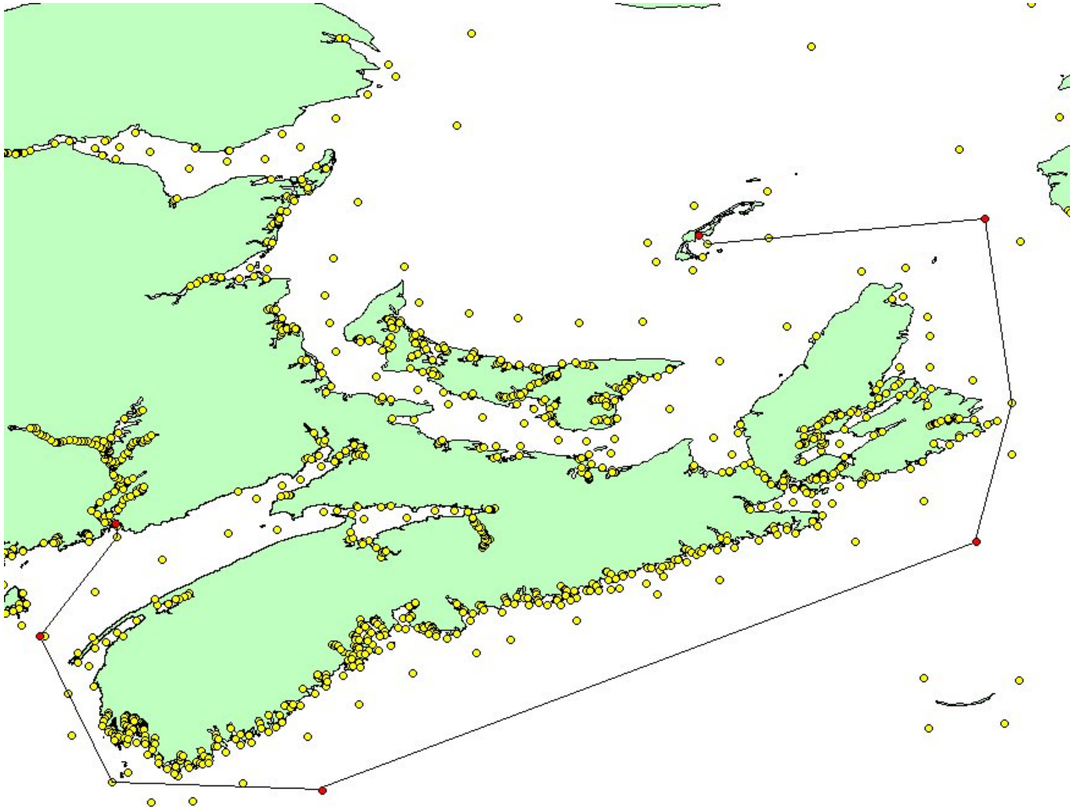
Direct Connection Results

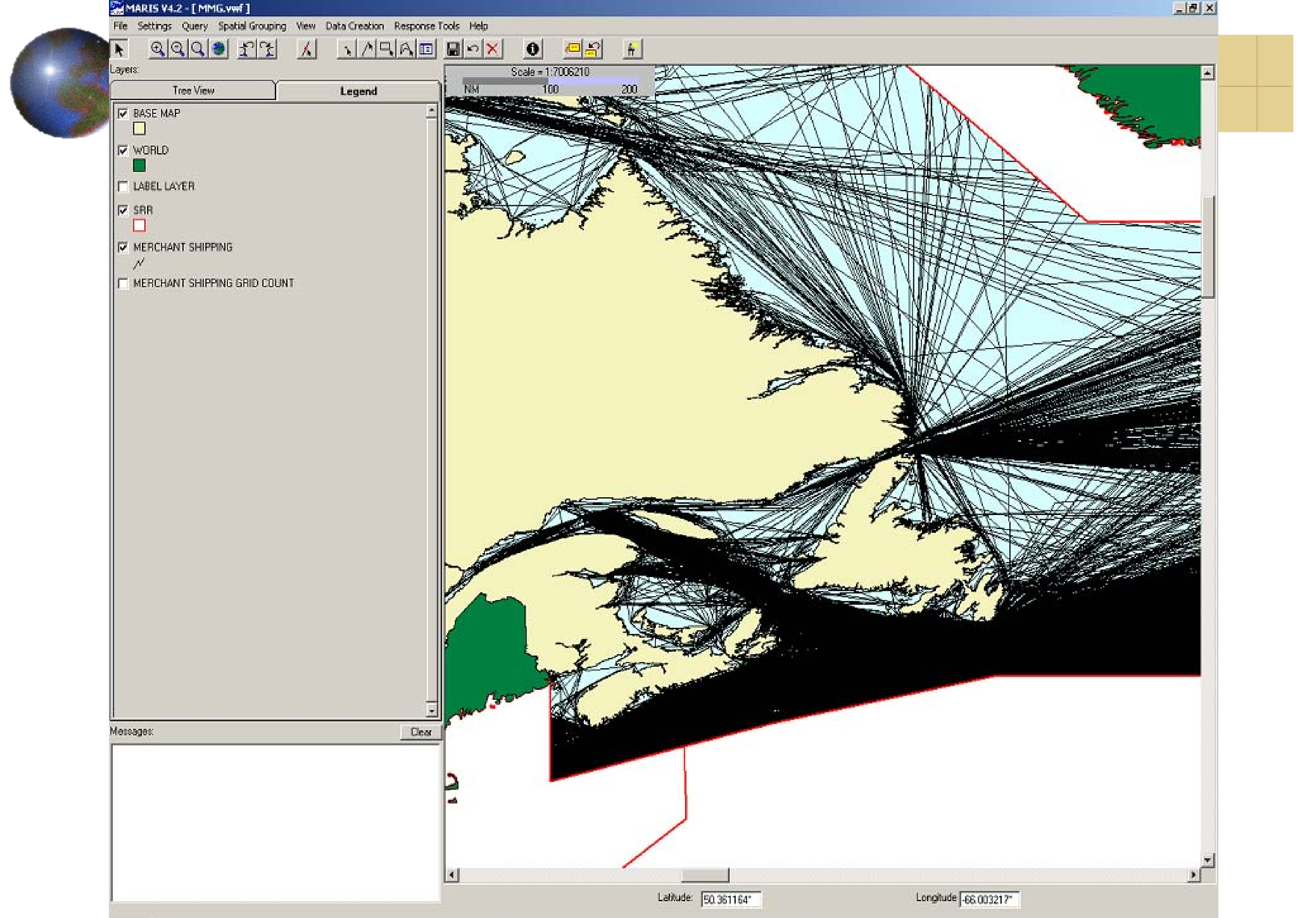


Node Network

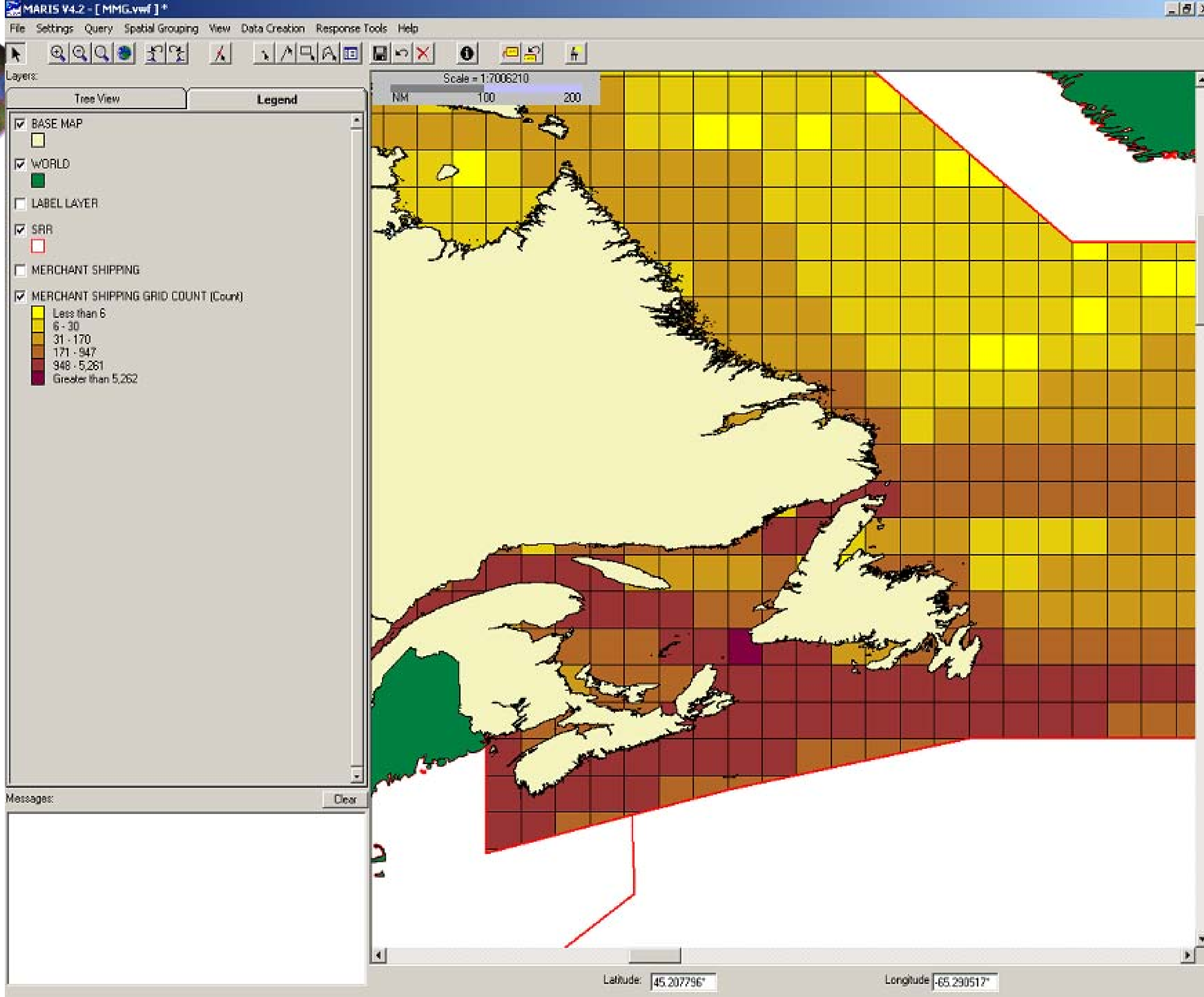


Land Avoidance



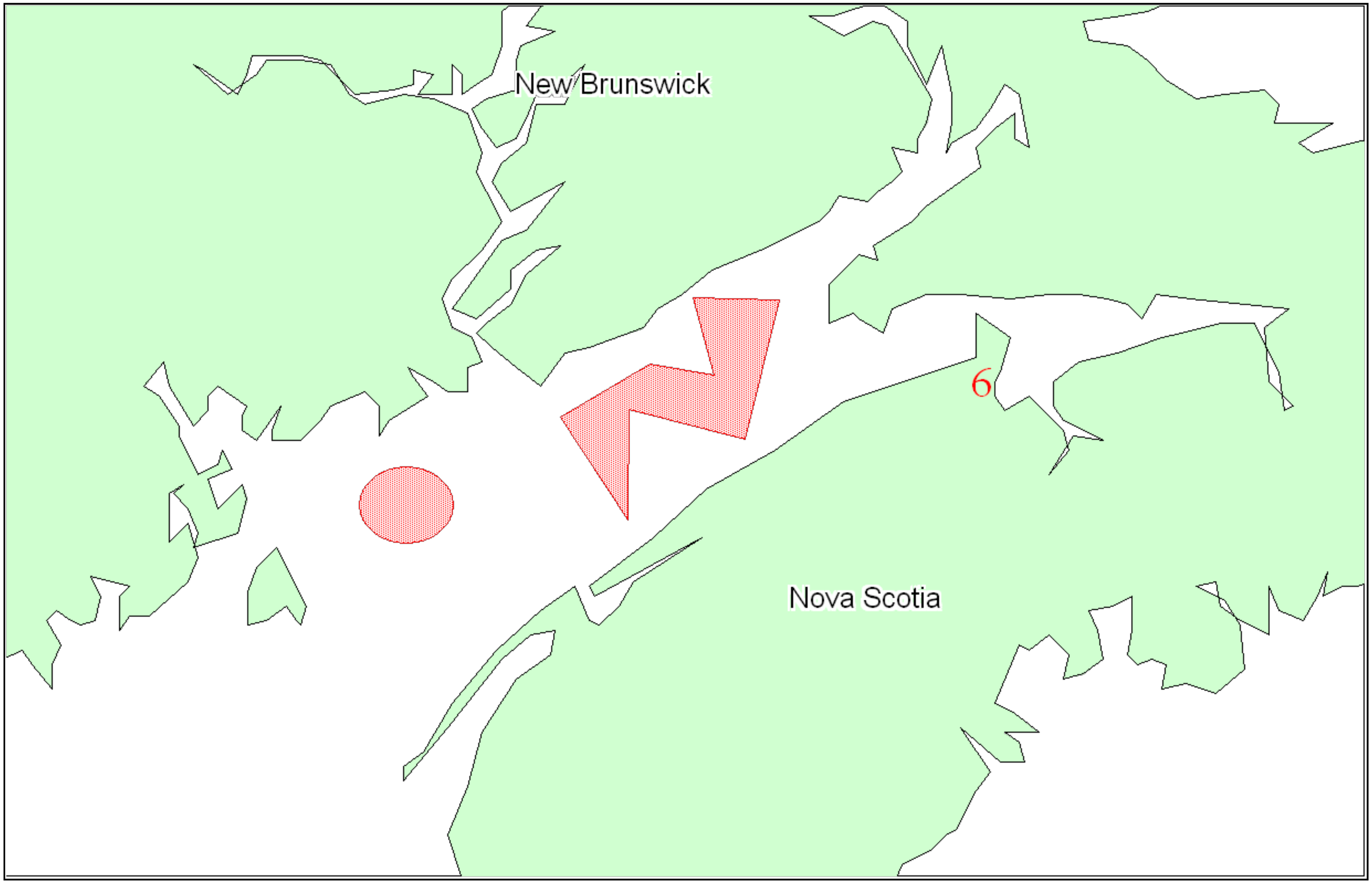


MARIS V4.2 - [MMG.vwf]
File Settings Query Spatial Grouping View Data Creation Response Tools Help
Layers:
Tree View
Legend
BASE MAP
WORLD
LABEL LAYER
SRR
MERCHANT SHIPPING
MERCHANT SHIPPING GRID COUNT
Messages:
Clear
Scale = 1:7006210
NM 100 200
Latitude: 50.361164°
Longitude -66.003217°

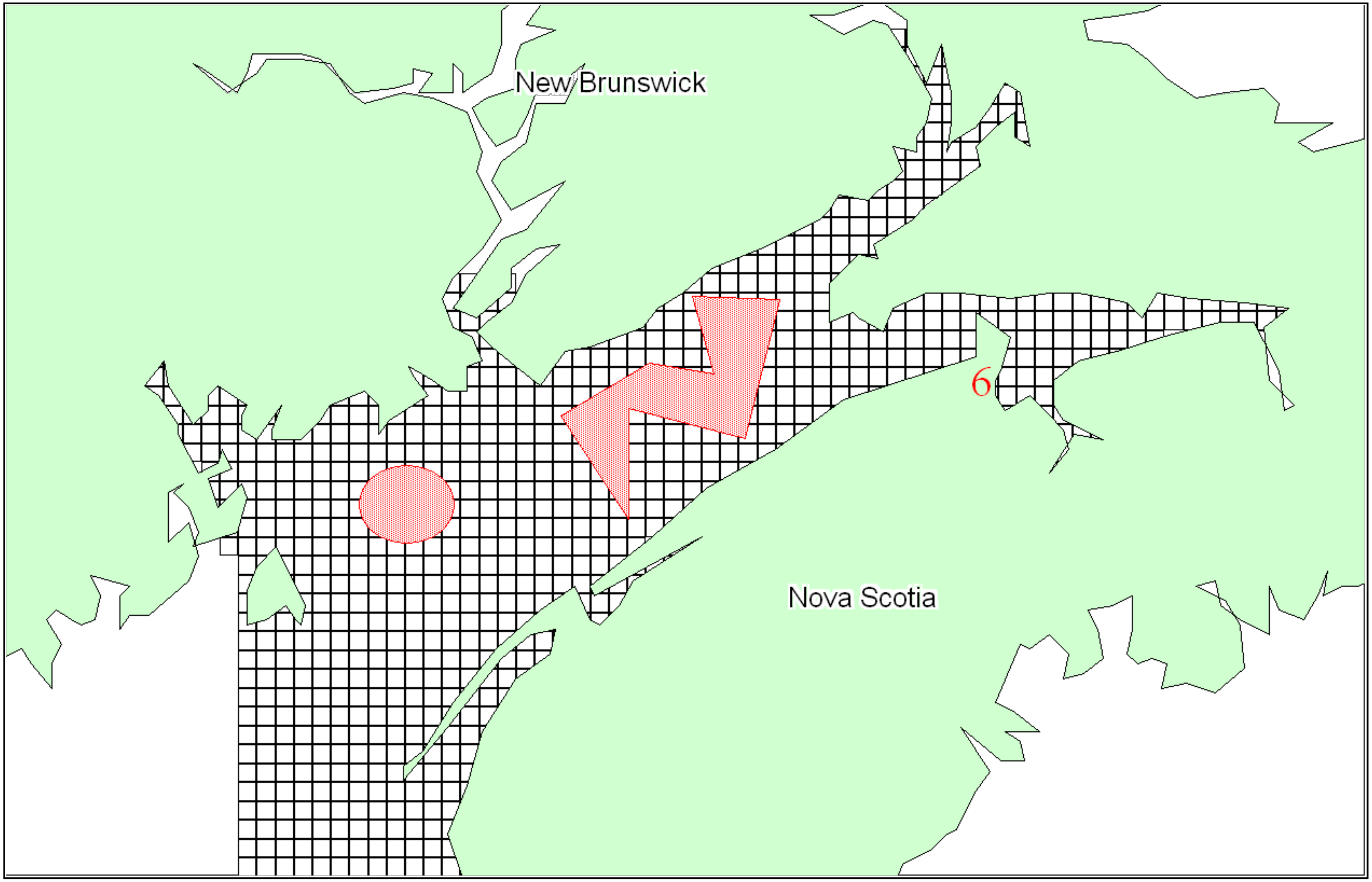

MARIS V4.2 - [MMG.vwf] *
File
Settings
Query
Spatial Grouping
View
Data Creation
Response Tools
Help
Layers:
Tree View
Legend
BASE MAP
WORLD
LABEL LAYER
SRR
MERCHANT SHIPPING
MERCHANT SHIPPING GRID COUNT (Count)
Less than 6
6 - 30
31 - 170
171 - 947
948 - 5,261
Greater than 5,262
Messages:
Clear
Scale = 1:7006210
NM
100
200
Latitude:
45.207796°
Longitude
-65.290517°

Dispersion Algorithm

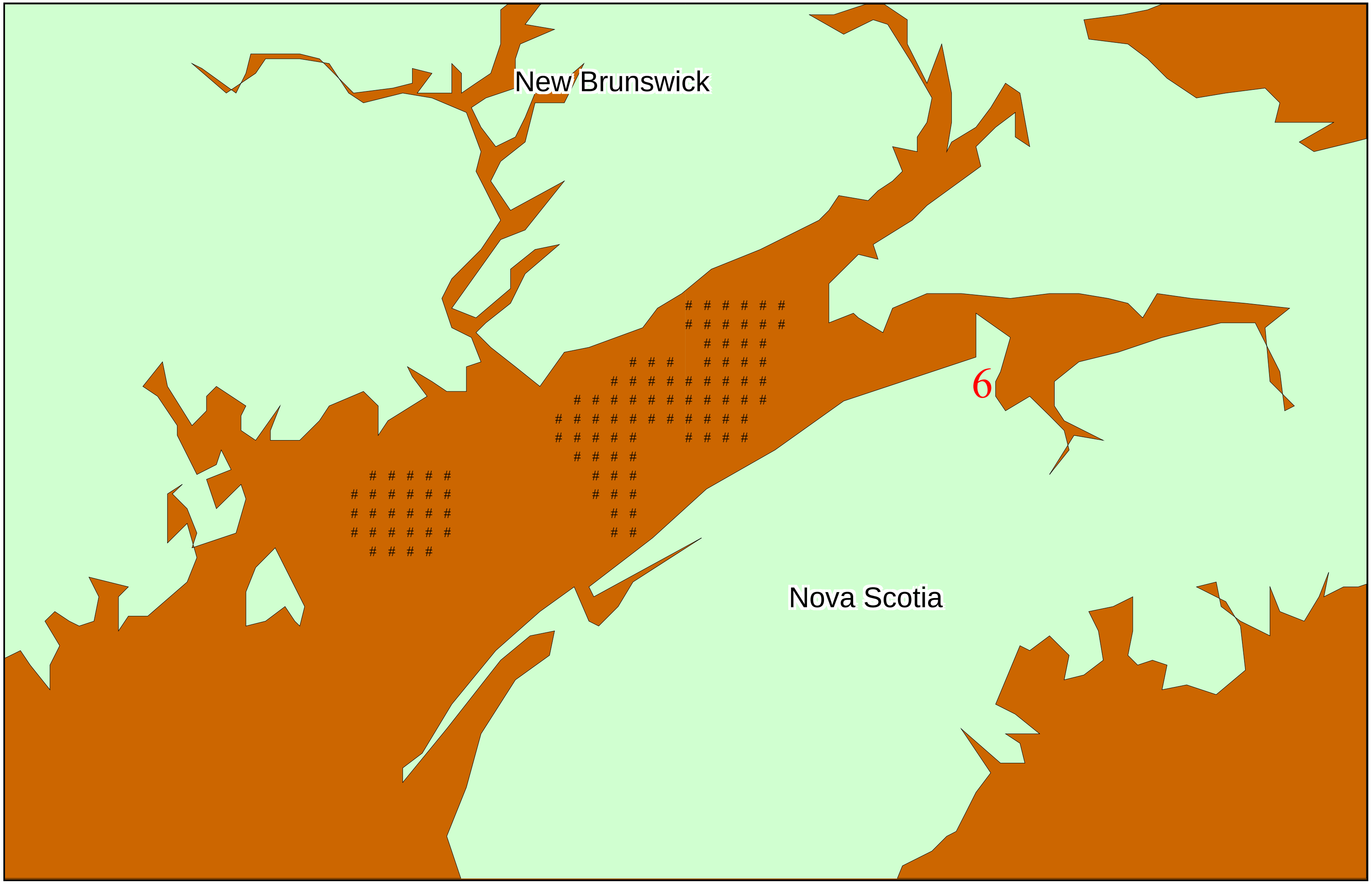
Trips to feasible region



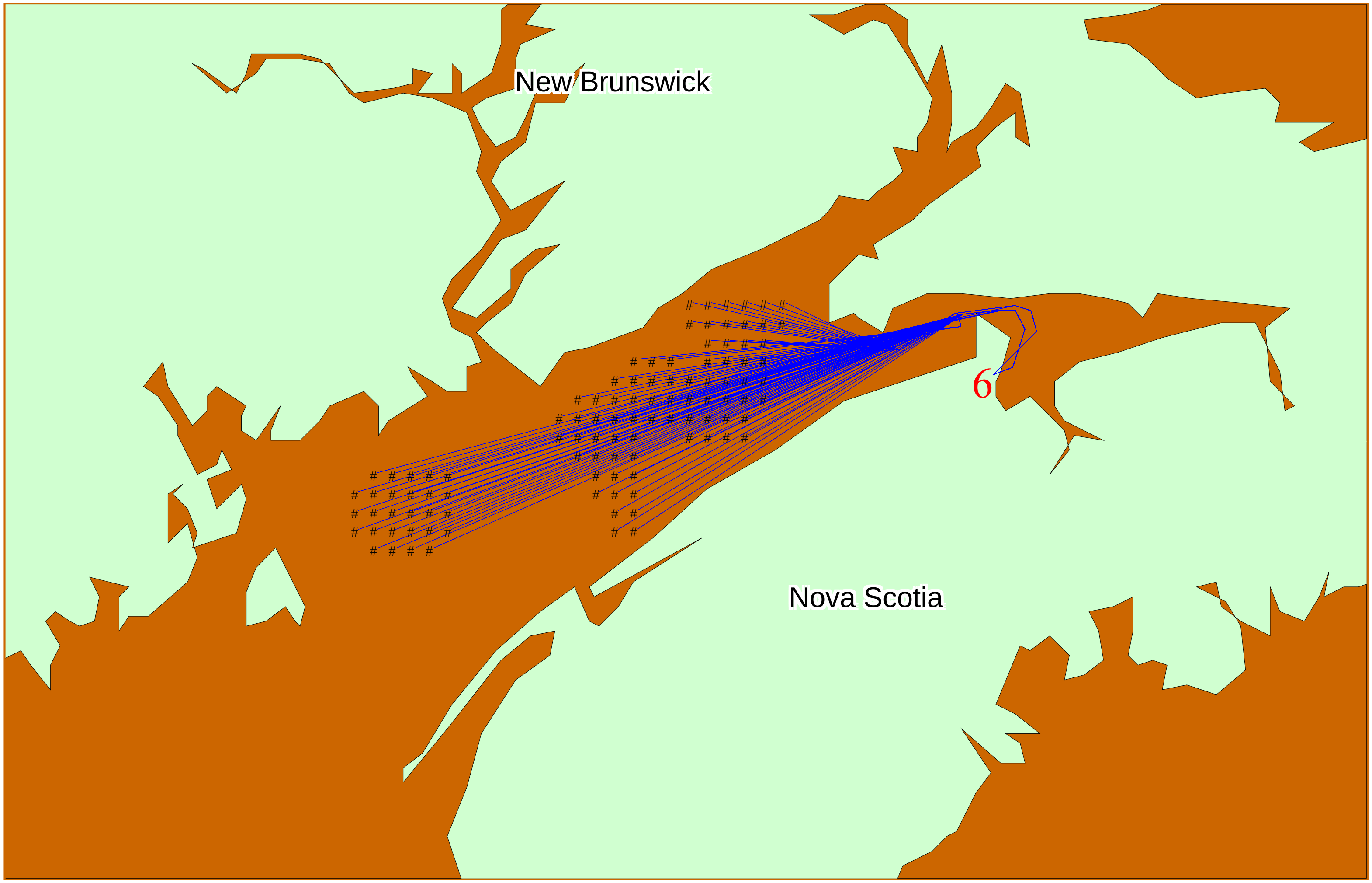
Grid the area



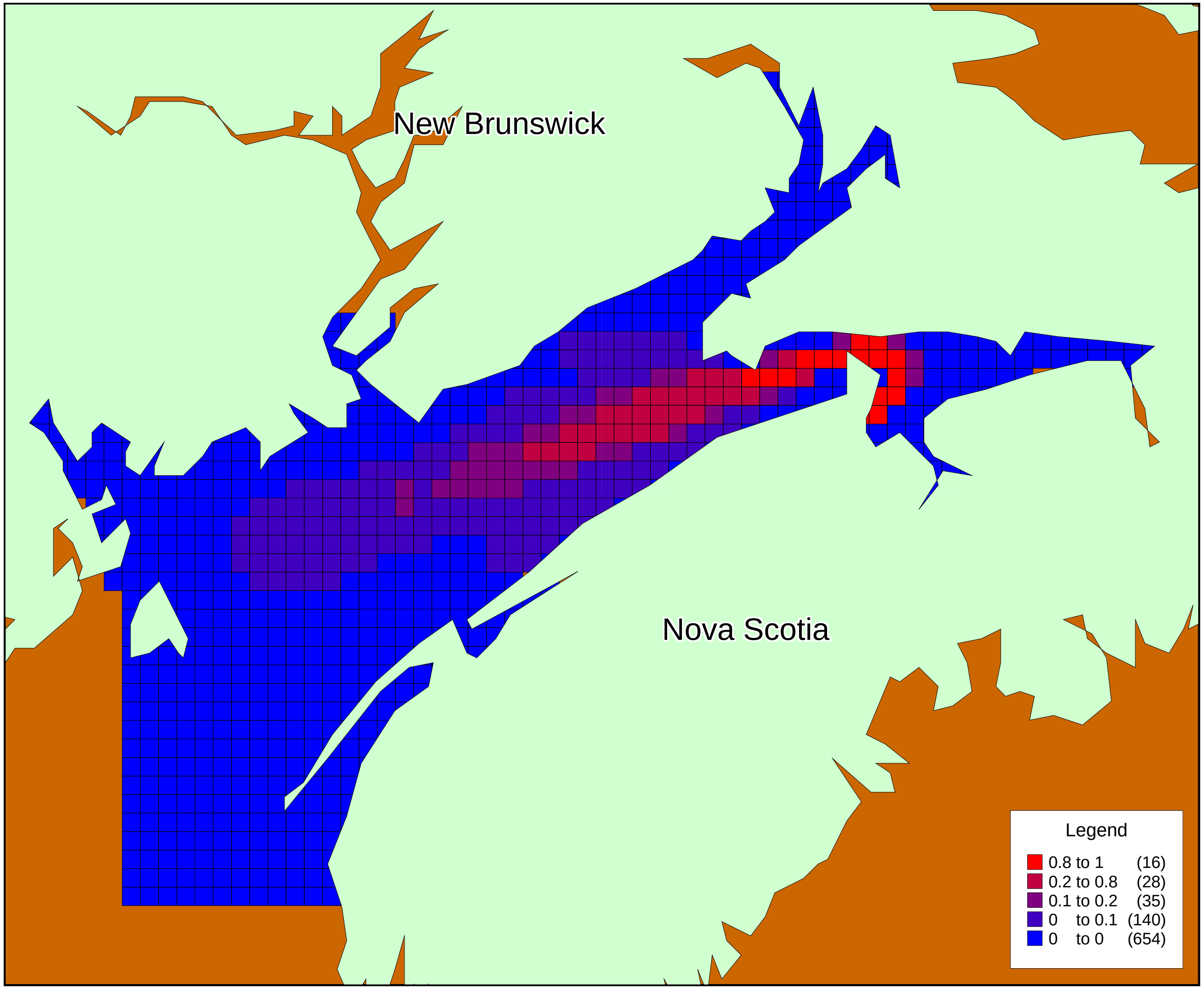
Define feasible destinations



Generate tracks



Grid results



Abnormal ship trajectories

Objectives

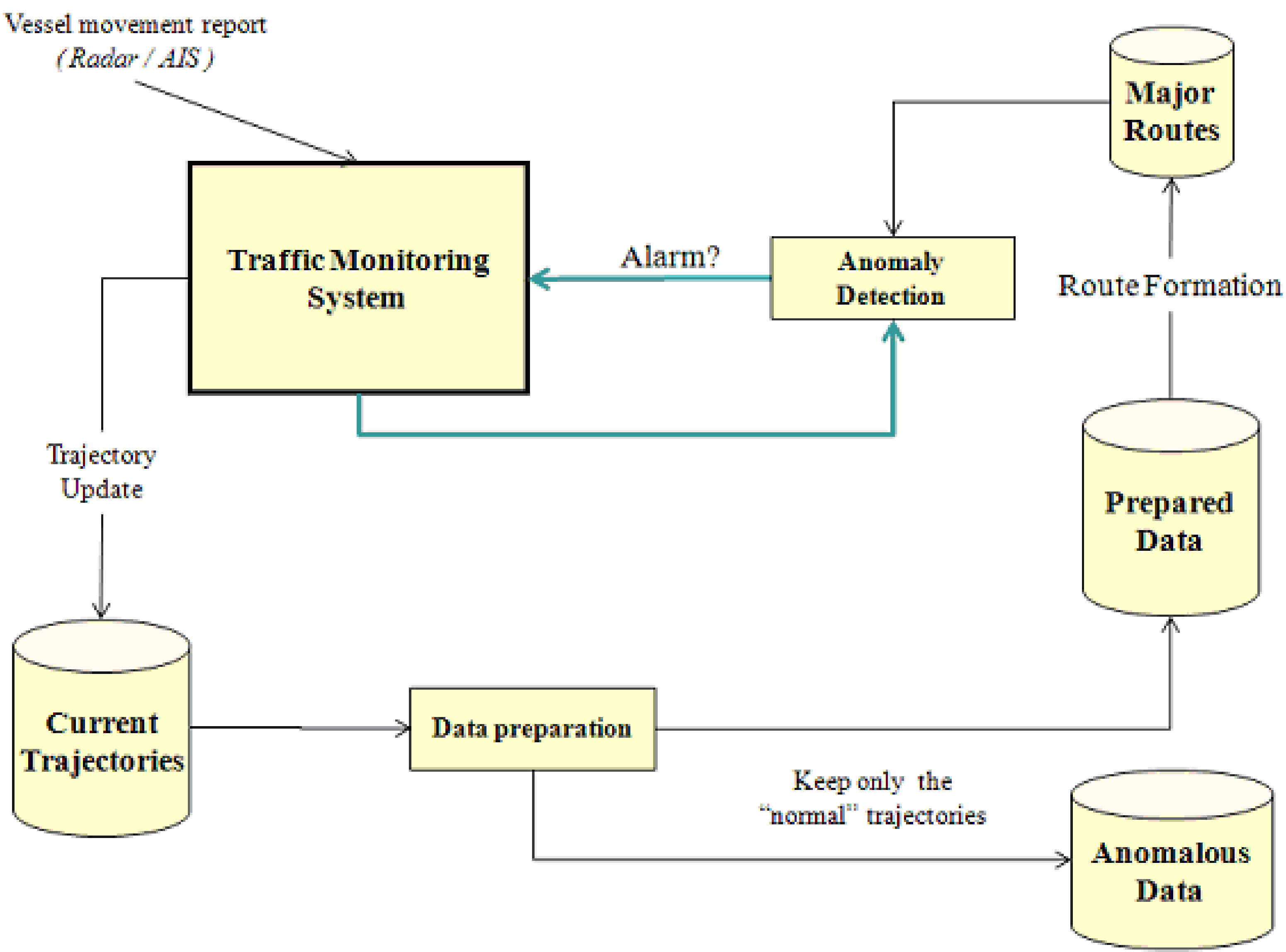
(1) To mathematically define commonly used shipping routes from historical shipping data.

(2) To detect unusual vessel trajectories based on norms of typical shipping routes using Statistical Process Control (SPC).

Line pattern definitions

- Line pattern analysis in this research involves the description of the relationship among different trajectories and the classification of similar **trajectories** into **routes** and **paths**.
- **Trajectory**: a line generated from a series of waypoints by an individual moving object.
- **Route**: a frequently used track followed by moving objects from the time when they enter the study area until the time they exit.
- **Path**: a feature grouping common sections of routes.

Traffic Monitoring System



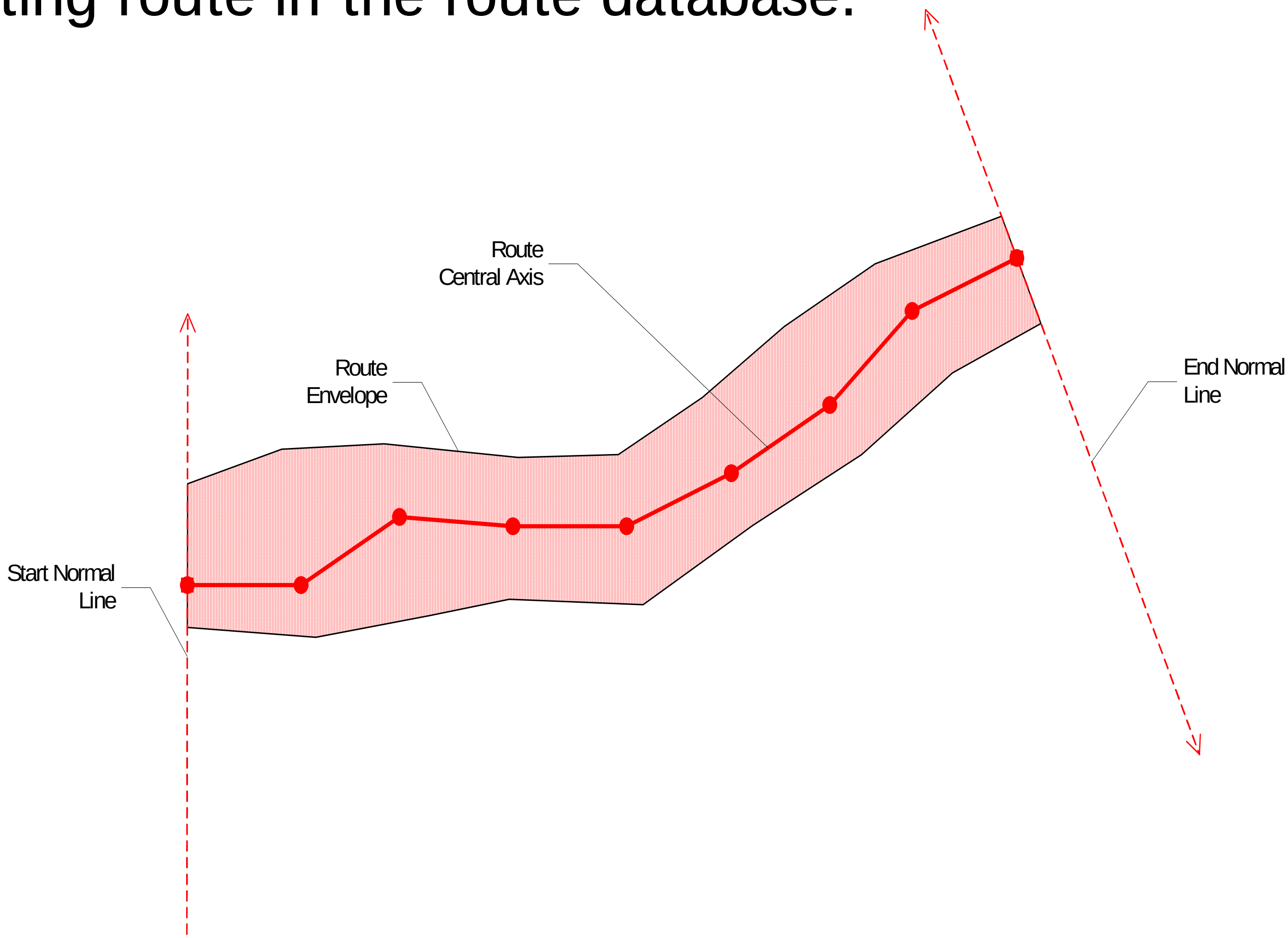
Geometric Algorithm Review

Three main steps:

- Route-trajectory **matching**;
- Route-trajectory **updating**;
- Route-route **merging**.
- Each route node here is characterized by a vector format with coordinates $\vec{x}_i = [x_i, y_i]$.
- Attributes of each route: central axis, left and right boundaries (also called "envelopes"), normal vector and weight.
- Resample distance – equally spaced nodes are obtained on each trajectory and route.

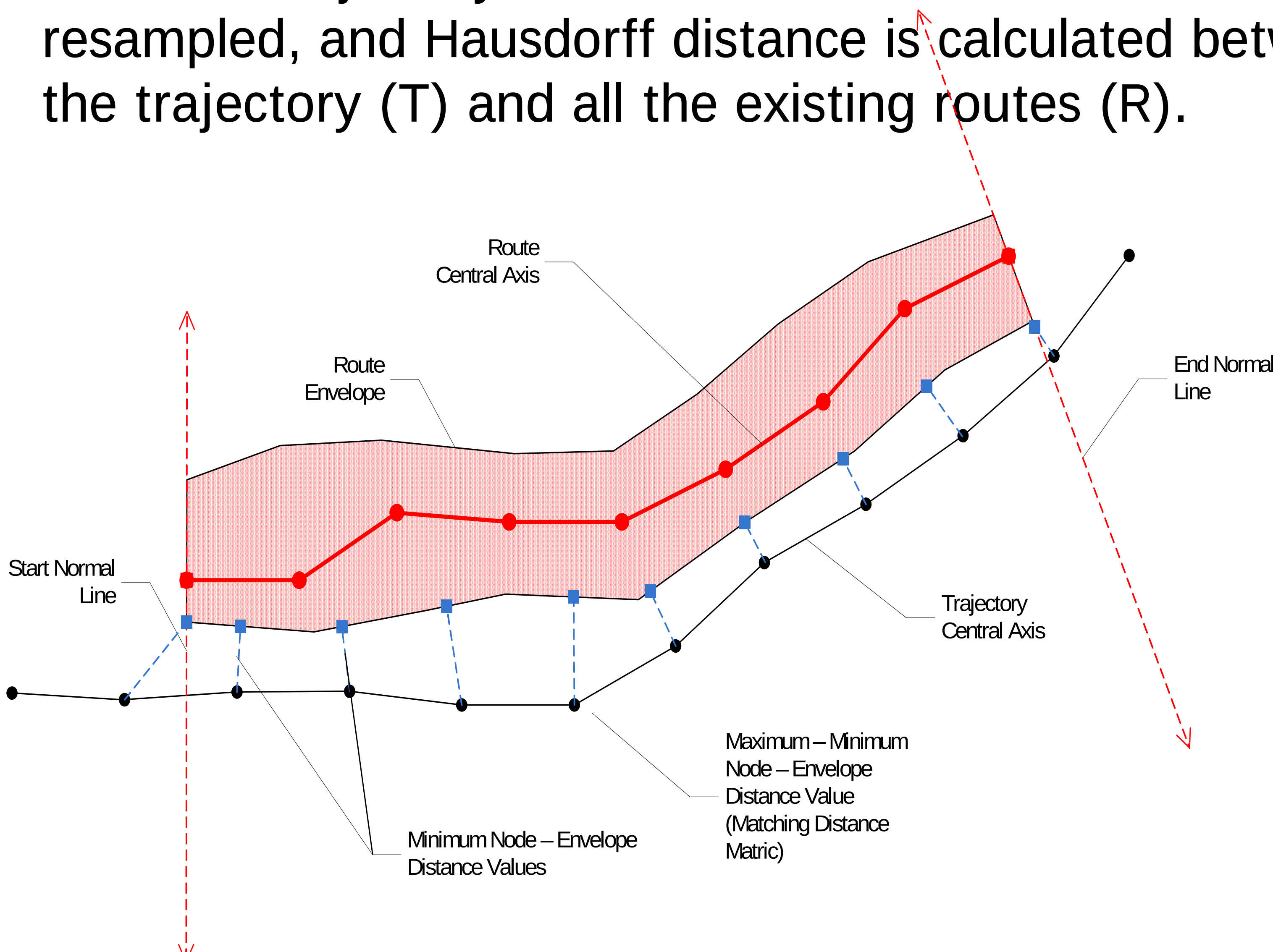
Route-Trajectory Matching

- Label each trajectory with a unique identifier and resample it. The first trajectory is initialized as the first existing route in the route database.



Route-Trajectory Matching

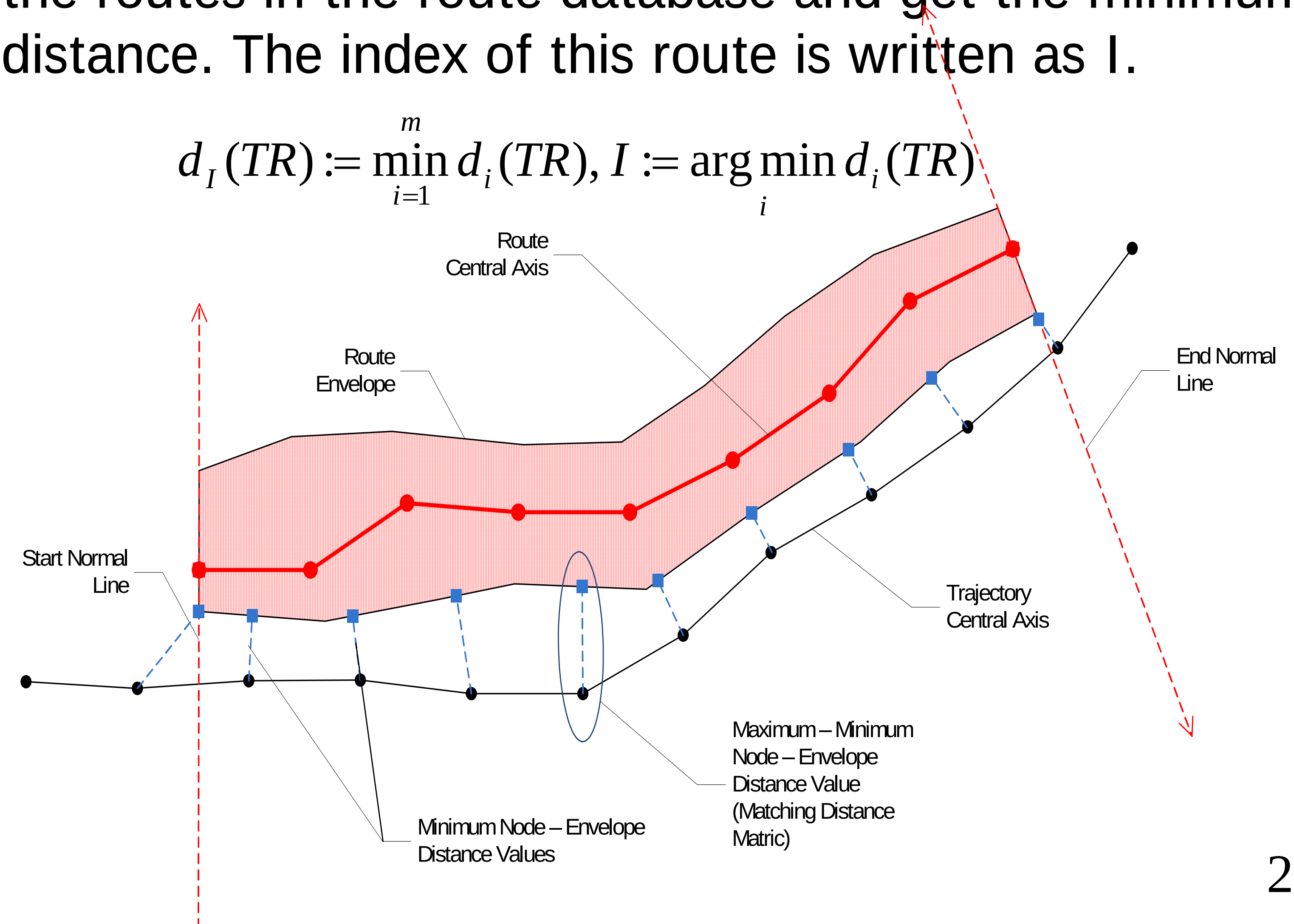
- A second trajectory is taken from the database, resampled, and Hausdorff distance is calculated between the trajectory (T) and all the existing routes (R).



Route-Trajectory Matching

- Compare the distances $d_i(TR)$ from the trajectory to all the routes in the route database and get the minimum distance. The index of this route is written as I.

$$d_I(TR) := \min_{i=1}^{m} d_i(TR),\ I := \arg\min_{i} d_i(TR)$$



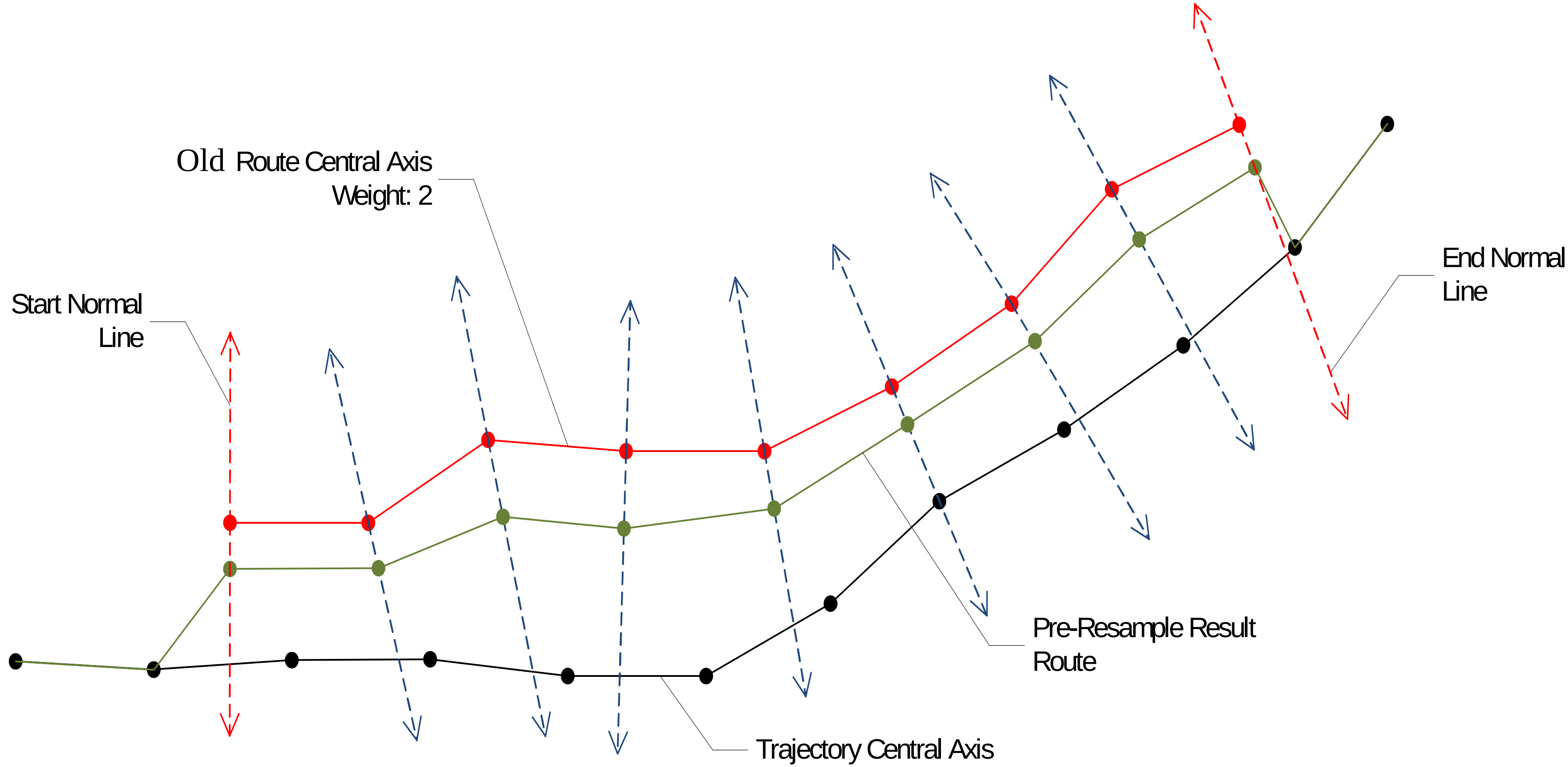
Route-Trajectory Matching

- If d_l(TR) is greater than a given threshold (‘*trajectory - assimilate-threshold*’), the trajectory will be created as a new route.

- Otherwise, a match between the trajectory and route l is detected. Go to 2nd step, route updating.

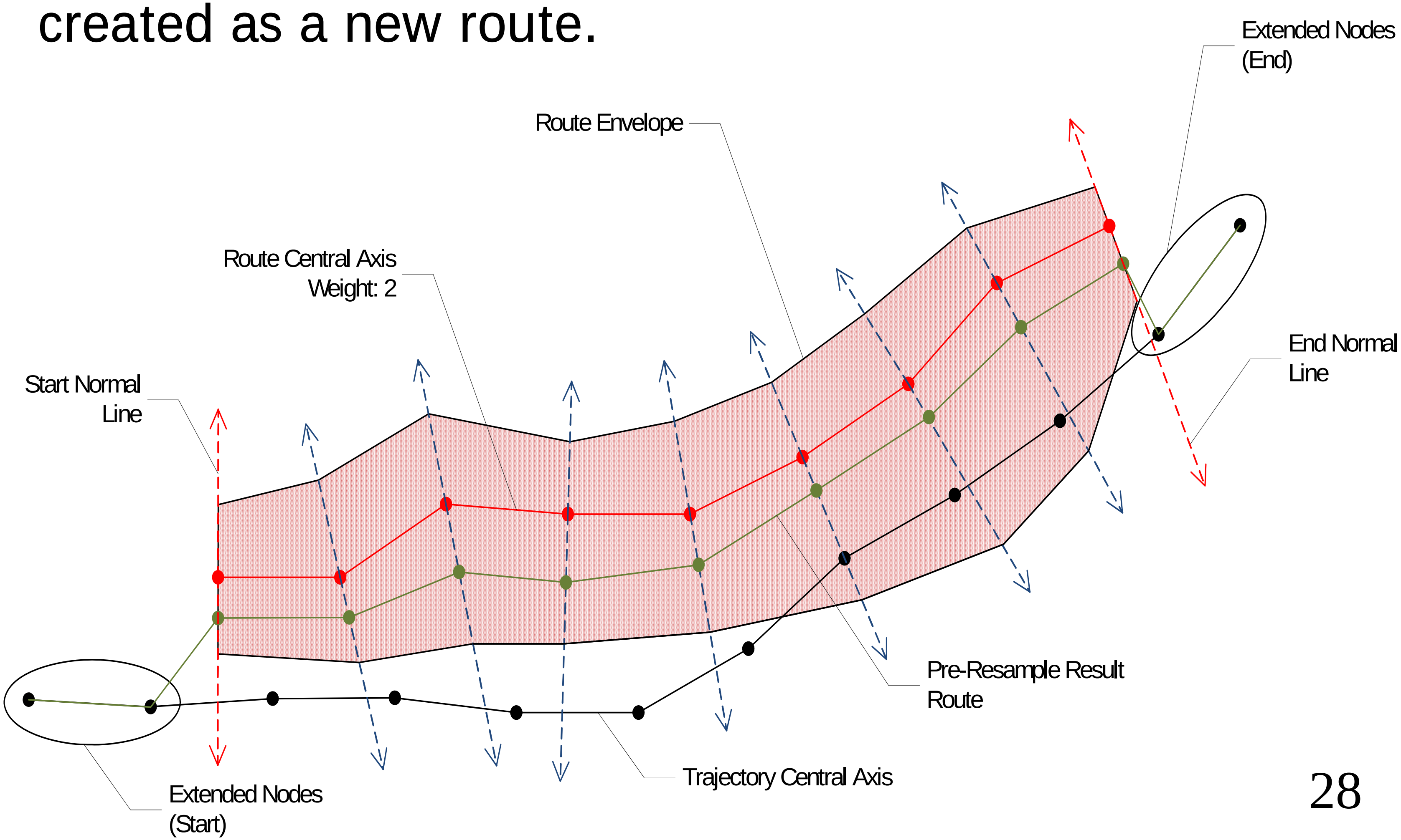
Route-Trajectory Updating

- Central axis node: $\overrightarrow{X_{Ik}} := \frac{w_I}{w_I + 1} \cdot \overrightarrow{x_{Ik}} + \frac{1}{w_I + 1} \cdot \overrightarrow{x_{Ikt}}$



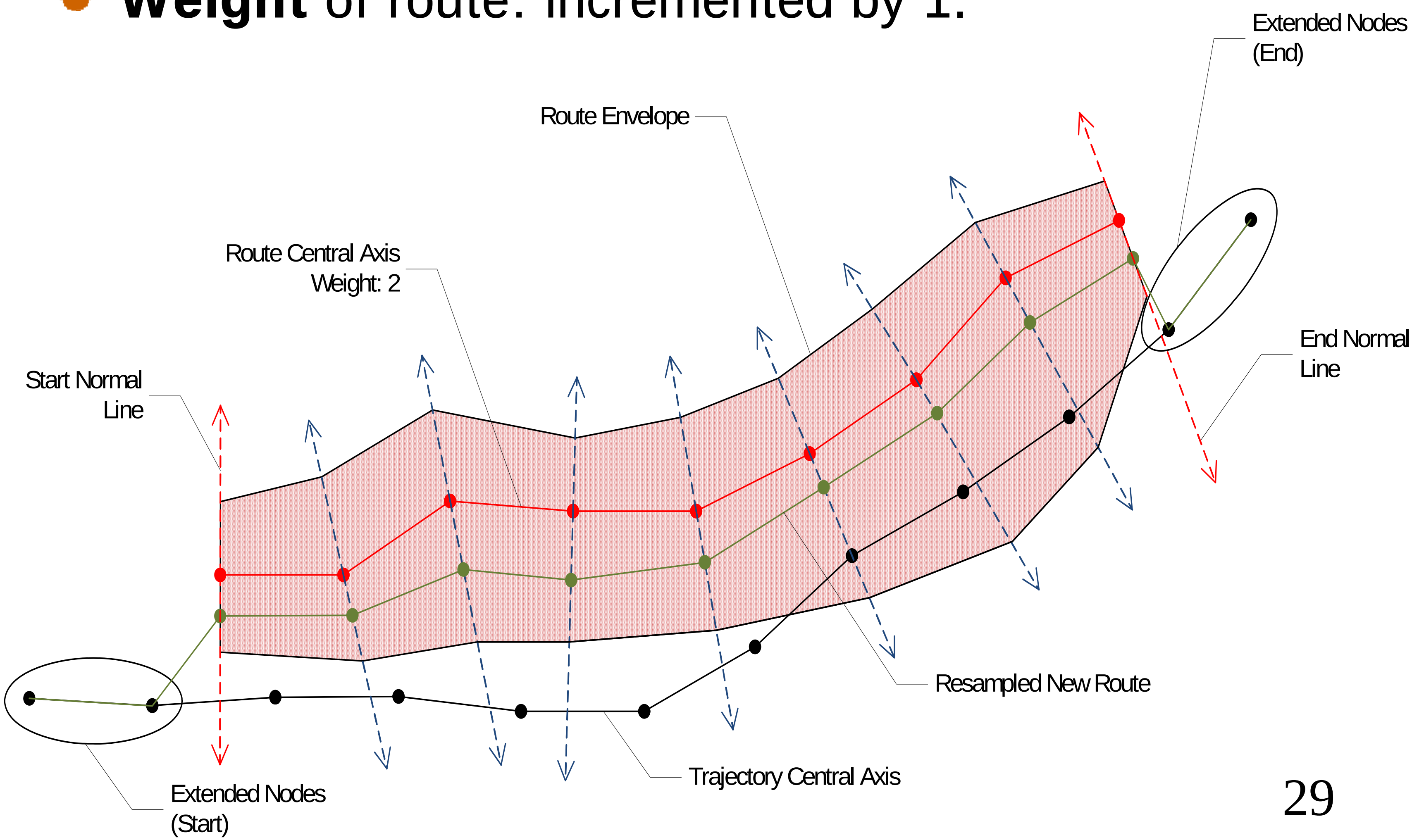
Route-Trajectory Updating

- Route **extension**: The longer part of the trajectory is used to extend the route I. If the extension length is more than half of the length of the entire route, the trajectory is created as a new route.



Route-Trajectory Updating

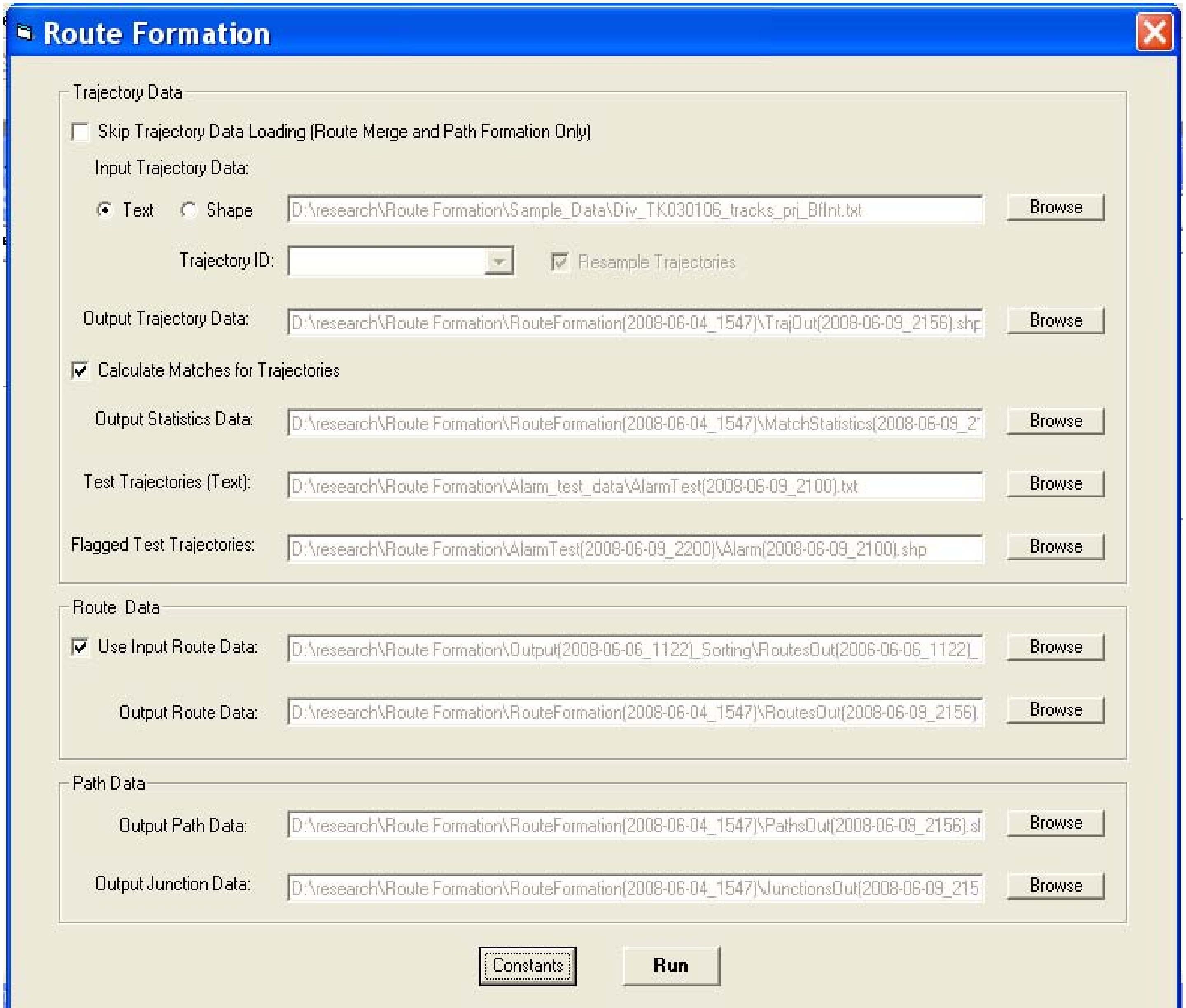
- **Boundaries**: The farthest intersections on both sides are the new left and right boundary nodes.
- **Weight** of route: incremented by 1.



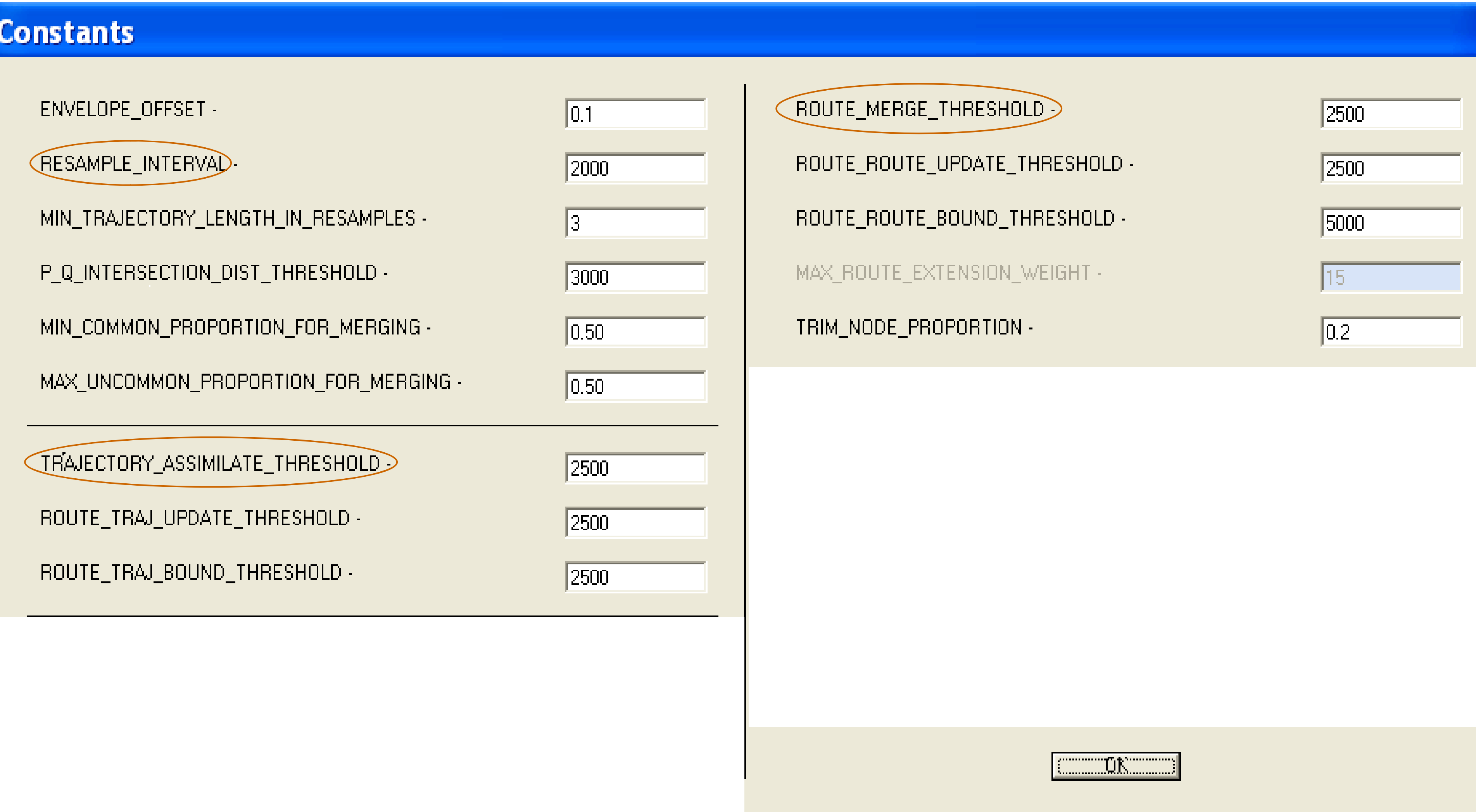
Route-Route Merging

- Sort all the routes by descending order of weights.
- Step 1: Start from the route with the highest weight (R_h). Calculate the Hausdorff distance between R_h and the route with the smallest weight in the database (R_s).
- Step 2: If the distance is less than a predefined threshold ('*route-merge-threshold*'), then the secondary route is qualified to be merged into the main route R_h. Otherwise, they are not merged.
- Step 3: Go to the next lower weighted route and repeat the above two steps until all the routes in the database are compared with R_h.
- Step 4: Go to the next higher weighted route and start a new loop (the above 3 steps), until all the routes (except the least weighted route) are used as R_h to be compared.

User file control



User parameter control (metres)



Constants

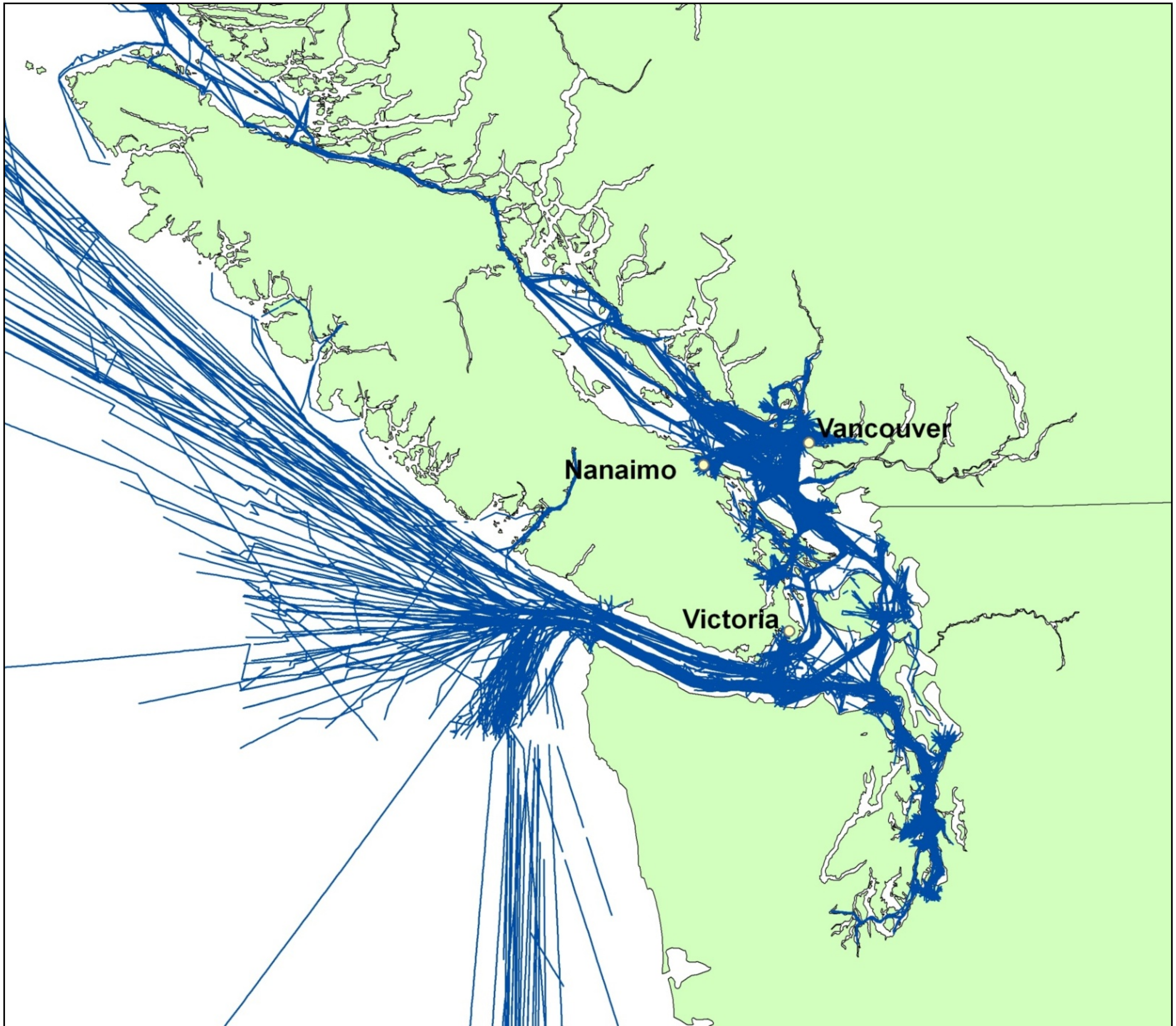
ENVELOPE_OFFSET -	0.1
RESAMPLE_INTERVAL -	2000
MIN_TRAJECTORY_LENGTH_IN_RESAMPLES -	3
P_Q_INTERSECTION_DIST_THRESHOLD -	3000
MIN_COMMON_PROPORTION_FOR_MERGING -	0.50
MAX_UNCOMMON_PROPORTION_FOR_MERGING -	0.50
TRAJECTORY_ASSIMILATE_THRESHOLD -	2500
ROUTE_TRAJ_UPDATE_THRESHOLD -	2500
ROUTE_TRAJ_BOUND_THRESHOLD -	2500
ROUTE_MERGE_THRESHOLD -	2500
ROUTE_ROUTE_UPDATE_THRESHOLD -	2500
ROUTE_ROUTE_BOUND_THRESHOLD -	5000
MAX_ROUTE_EXTENSION_WEIGHT -	15
TRIM_NODE_PROPORTION -	0.2

OK

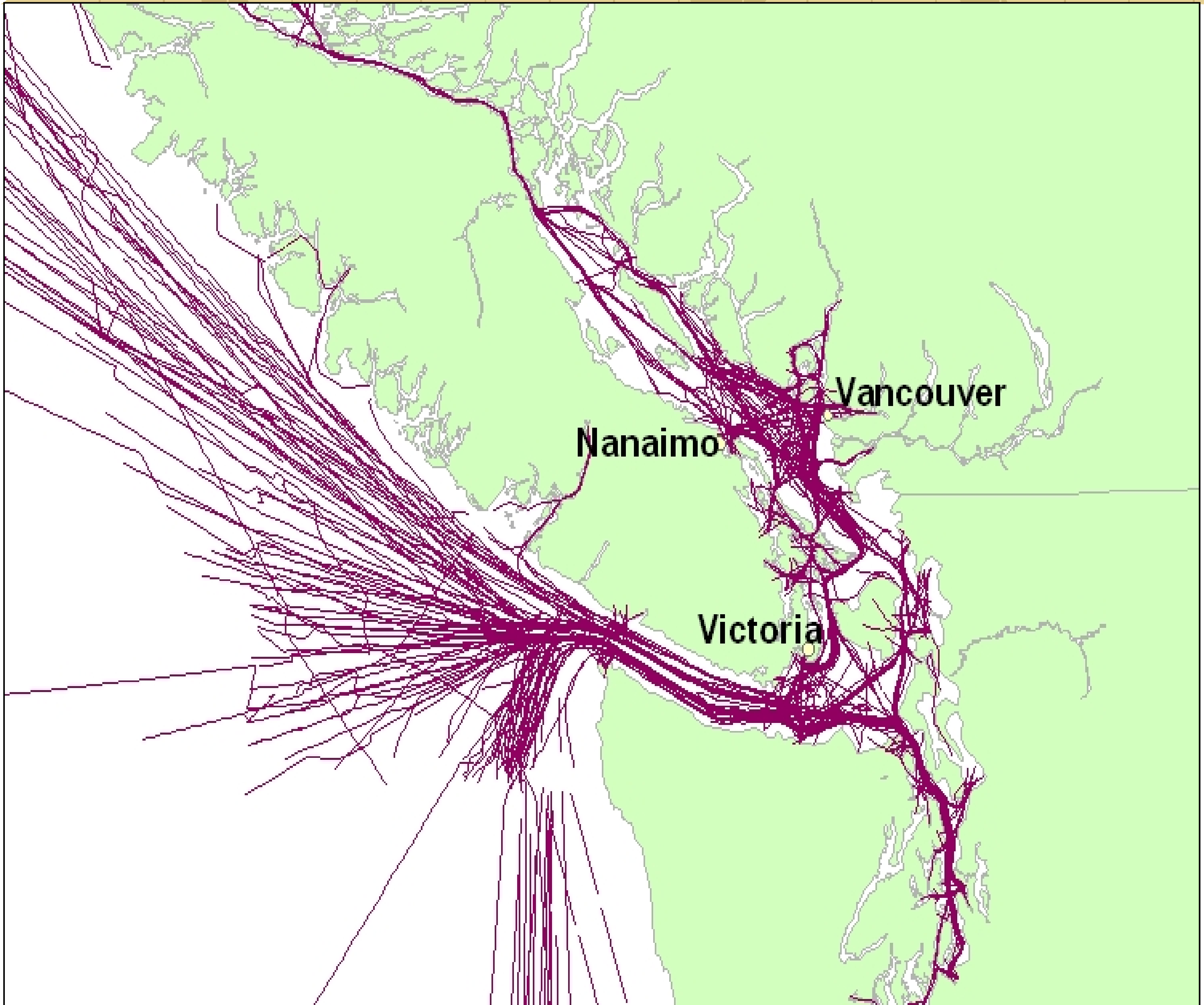
Implementation of the algorithm

- Data source: VTOSS database (Vessel Traffic Operations and Support Systems)
- Study area: the ships that approach within 50km of the British Columbia (BC) shoreline.
- Study period: Fourteen daily datasets from Jan. 1st to 14th, 2003; three weekly datasets – Jan. 1st to 7th, 8th to 14th, and Aug. 6th to 12th, 2003.
- The following figures are displayed for a one-week traffic data (from Jan. 1st to 7th, 2003).

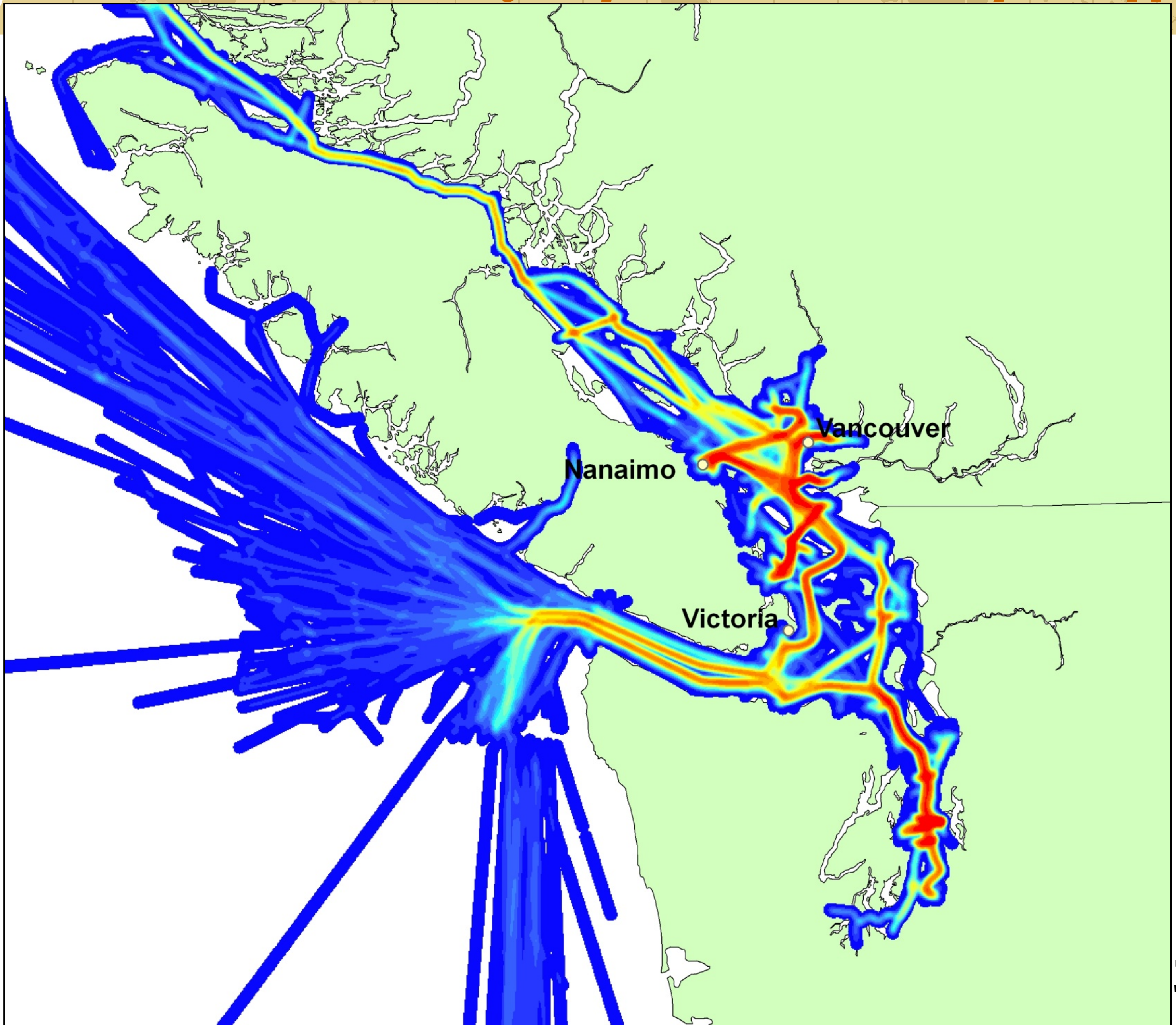
Historical trajectories



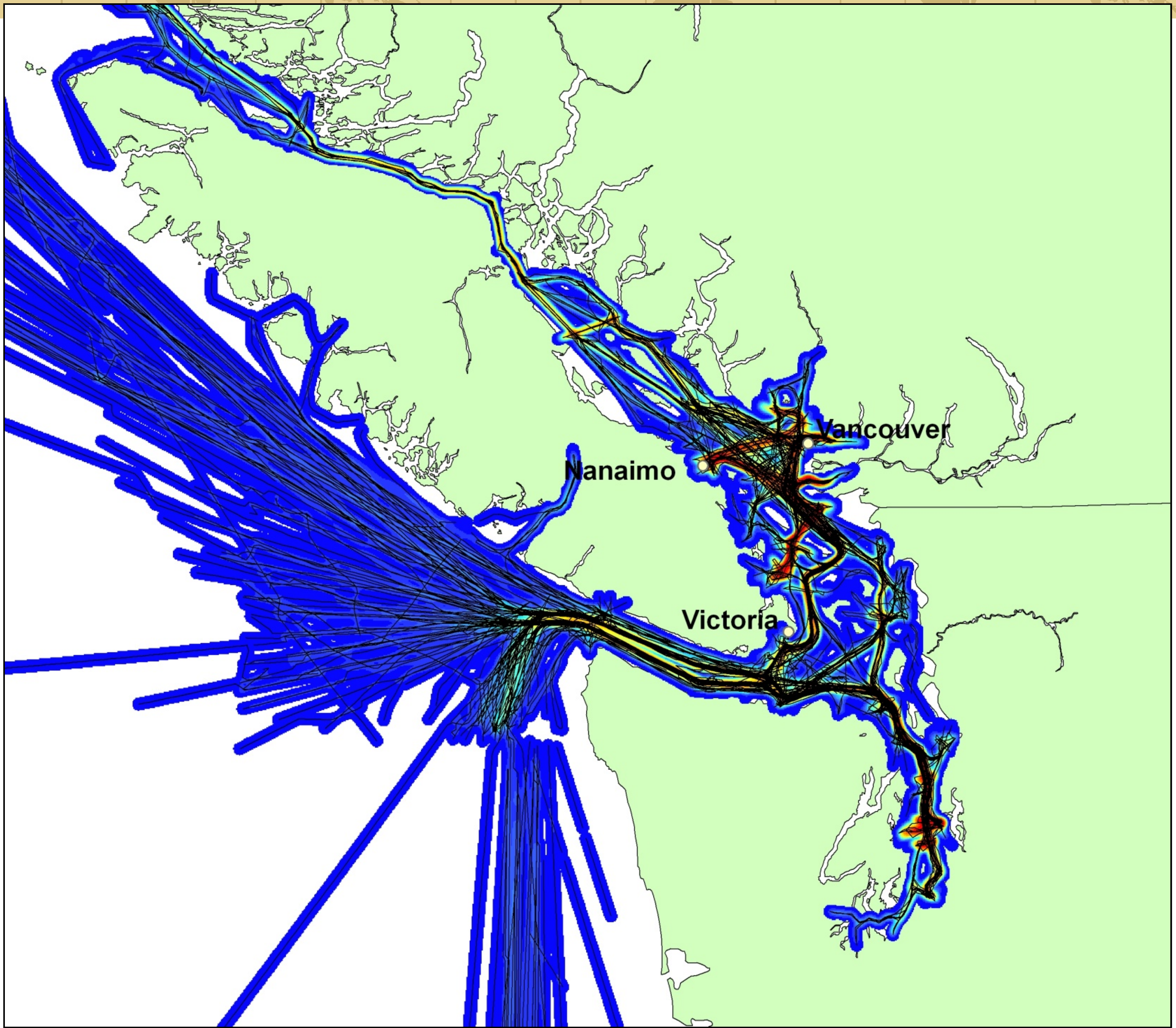
(resample_interval = 2500, trajectory_assimilate_threshold = 3000, route_merge_threshold = 3000)



Kernel Density of one week of traffic



Kernel Density vs Generated Routes



SPC Methodology Review

- Cumulative sum (CUSUM) control chart is applied in maritime traffic to monitor the process (i.e. movement) of any new individual trajectory.
- Let x_i be the i^{th} observation on the process. μ_0 is the target value for the quality characteristic *x*.
- The tabular CUSUM works by accumulating deviations from μ_0 that are above target with one statistic C^+ and accumulating deviations from μ_0 that are below target with C^-.

$$C_i^+ = \max[0, x_i - (\mu_0 + K) + C_{i-1}^+]$$

$$C_i^- = \max[0, (\mu_0 - K) - x_i + C_{i-1}^-]$$

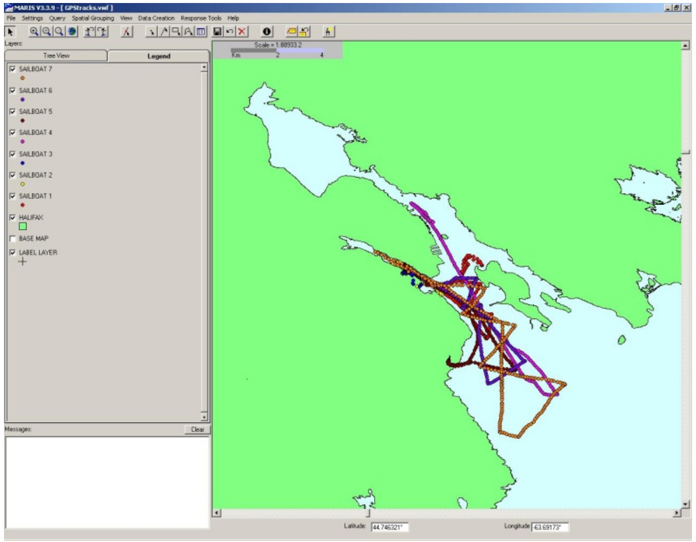
SPC Methodology Review

- x_i = the minimum distance from node *i* on the trajectory to the route envelope.
- The overall mean and standard deviation (unbiased) are calculated for the route across all incorporated trajectory nodes to route envelope distances.
- Only C^+ statistic is used, because distance is a positive variable.
- Two levels of alarm are set.

Alarm test of two individual trajectories

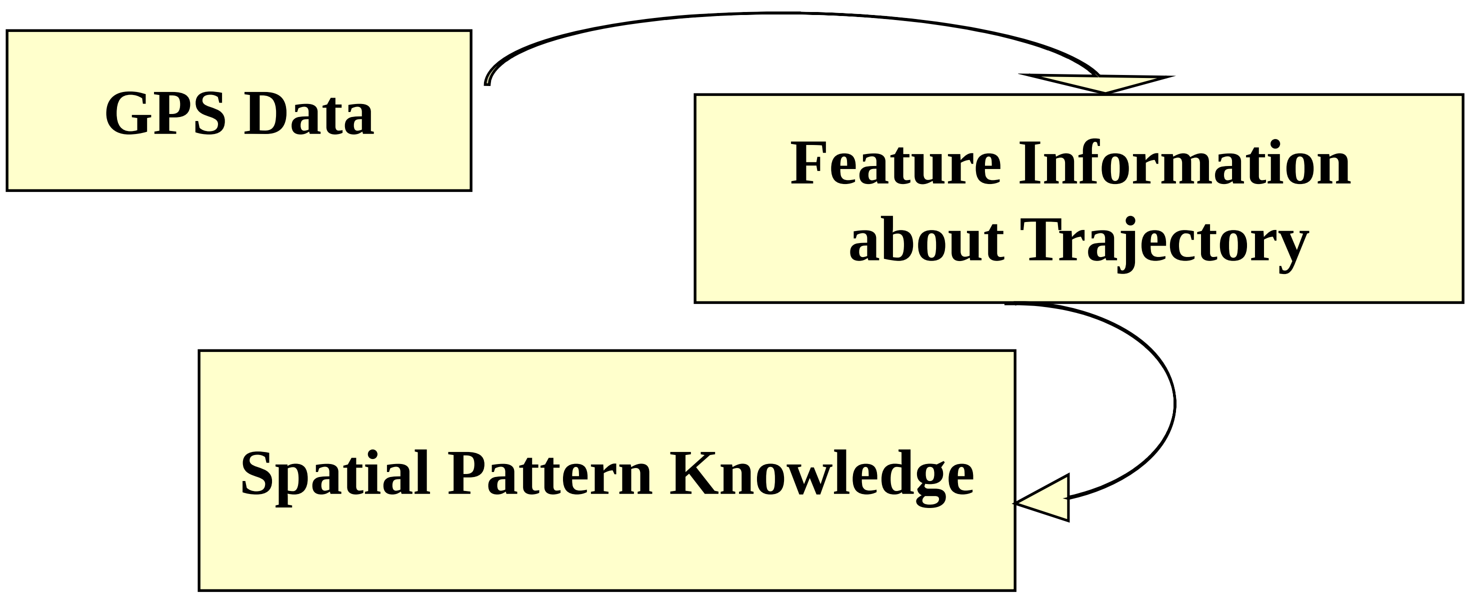
Discriminating recreational boat types

Characterizing of Recreational Boats based on GPS Trajectory Points



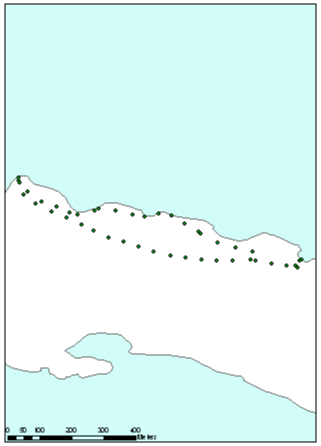
Objectives

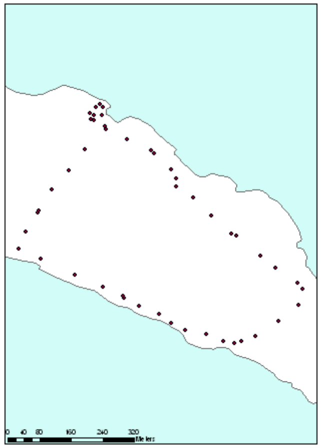
This Study examines Global Positioning System (GPS) trajectory data to characterize spatial patterns, providing insight into recreational boat movements and amount of exposure to risk.

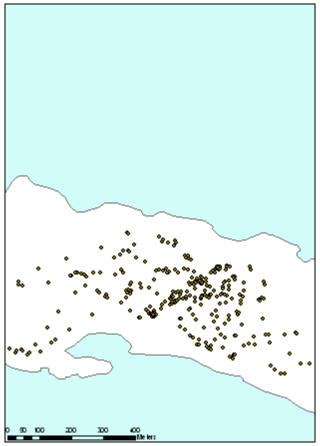


Spatial Pattern I

- **Patterns of Different Recreational Boat Types**



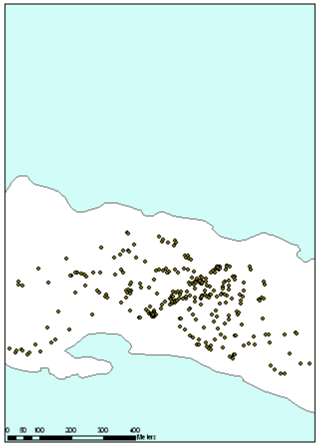




(a) c-0707-08-03 **(b) c-0721-11-04** **(c) s-0722-14-10**

Spatial Pattern II

- **Patterns of Different Recreational Boat Types at Different Locations**



(a) s-0722-14-10

(b) s-0824-16-03

Scope of Study

Sample size	NB	NS
Canoe	4	17
Kayak	8	21
Motorboat	15	10
Sailboat	5	47

Methodology

- Data Pre-cleaning
  - Points on land
  - Time gap
  - Resting/Stopping events
- Attributes Extracting/Calculation
  - Speed
  - Total distance
  - Bounding Box
  - Dedensified Trajectory
  - Distance from shore
- Discrimination & Classification

Attributes to Variates

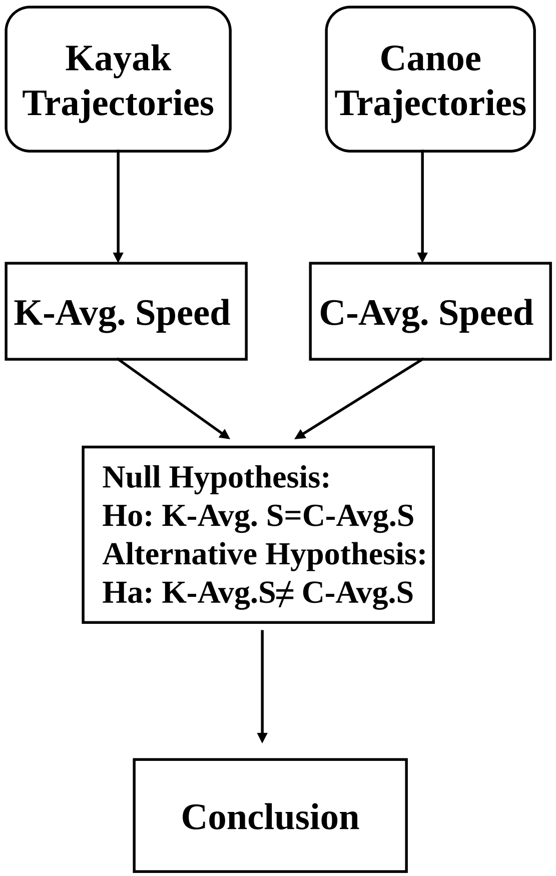
- **Speed**
- **Total distance**
- **Bounding Box**
- **Dedensified Trajectory**
- **Distance from shore**

→

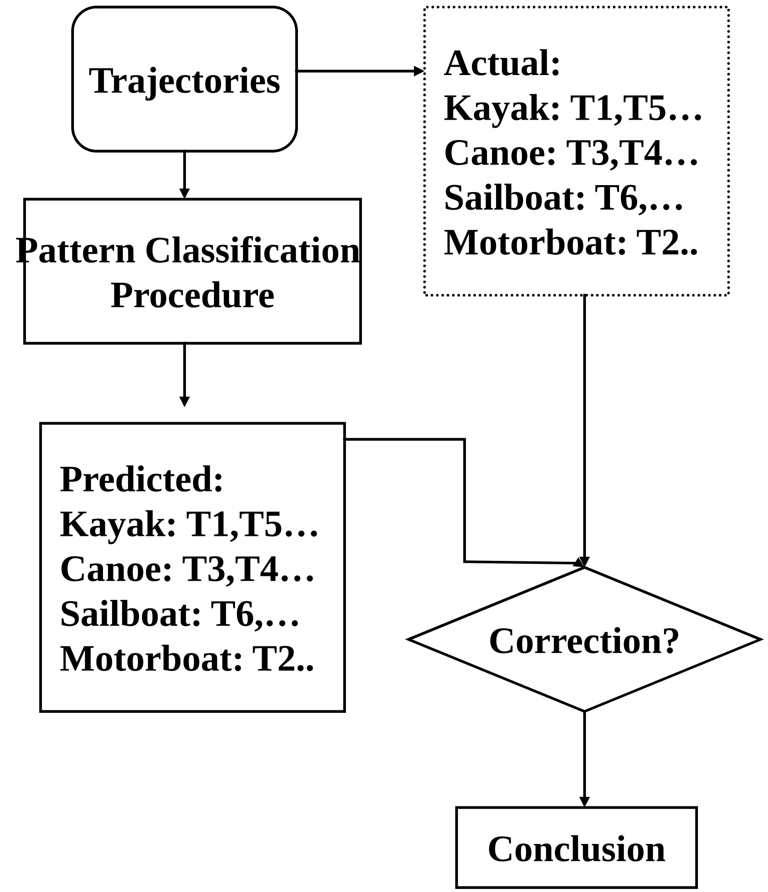
- **Max Speed**
- **Max $_{1/20}$ Speed**
- **Mean Speed**
- **Total Distance**
- **Complexity**
- **Aspect Ratio**
- **Mean Turning Angle**
- **Distance to Shore**

Pattern Analysis

➢ Discrimination



➢ Classification



Discriminant Analysis

Do different boat types have significant different attributes?

Attribute	NB	NS
TD		C, K S,M
Mean.S	C,K,S M	C,K M S
Max.S	C,K,S M	C,K M S
Max $_{1/20.}$S	C,K,S M	C,K M S
TAQ	C,K S	K S
TAH	C,K S	K S
C	M S	
DTS	C,K,M S	C,K S
AR	K S	C S

Discriminant Analysis

Do geography characteristics influence boating trajectories?

	Canoe	Kayak	Motorboat	Sailboat
TAH	X		X	X
TAQ	X	X	X	X
C	X	X	X	X
DTS	X	X		X
Max.S		X		X
Max $_{1/20}$.S		X		X
Mean.S		X	X	X
TD		X	X	X
AR	X		X	X

P=0.1

X: Not reject H_{0}. There is no significant difference between the two areas.

Multivariate Discriminant/Classification Analysis

- ***Problem*: Whether can we determine the type of recreational boat solely through the trajectory data without reference to the weather and operator factors, even geography?**
- ***Objectives:***
  - **Determining whether statistically significant differences exist between attributes of different boating trajectories of different boat types**
  - **Determining which of the independent variables account the most for the differences in the average score profiles of the four groups.**
  - **Establishing procedures for classifying boating trajectories on the basis of their score on a set of attributes.**

Estimation of the Simultaneous Discriminant Analysis

➢Standardized Canonical Discriminant Functions:

$$Z_1 = 0.047x_1 + 1.009x_2 + 0.585x_3 + 0.093x_4$$

$$Z_2 = 0.454x_1 - 0.584x_2 + 0.469x_3 + 0.671x_4$$

$$Z_3 = -0.744x_1 + 0.202x_2 + 0.732x_3 + 0.174x_4$$

➢Fisher's Linear Classification Functions:

$$Canoe = -6.941 - 5.223E-05X_1 + 3.301X_2 + 0.125X_3 + 5.894E-04X_4$$

$$Kayak = -5.629 - 1.516E-05X_1 + 3.251X_2 + 0.103X_3 + 3.209E-04X_4$$

$$Motorboat = -27.16 - 6.223E-05X_1 + 9.891X_2 + 0.191X_3 + 7.617E-05X_4$$

$$Sailboat = -13.74 + 3.057E-05X_1 + 5.156X_2 + 0.179X_3 + 1.498E-03X_4$$

x1=Total distance, x2=Mean speed, x3=mean turning angle, x4=farthest distance to shore.

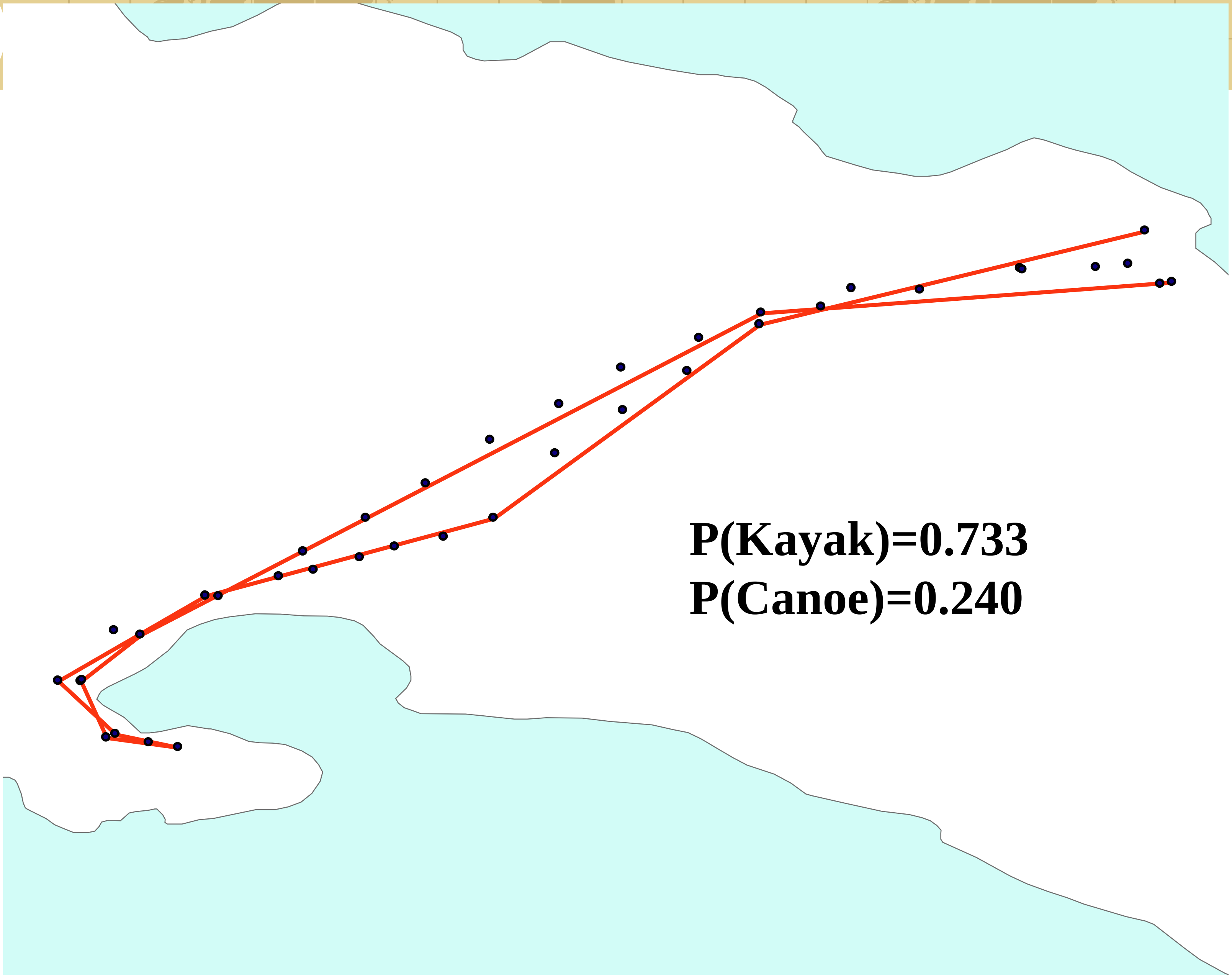
Classification Results

			Predicted Group Membership				Total
		TYPE	canoe	kayak	motorboat	sailboat	
Original	Count	canoe	15	2	0	0	17
		kayak	3	18	0	0	21
		motorboat	1	0	8	1	10
		sailboat	0	0	0	47	47
	%	canoe	88.2	11.8	.0	.0	100.0
		kayak	14.3	85.7	.0	.0	100.0
		motorboat	10.0	.0	80.0	10.0	100.0
		sailboat	.0	.0	.0	100.0	100.0

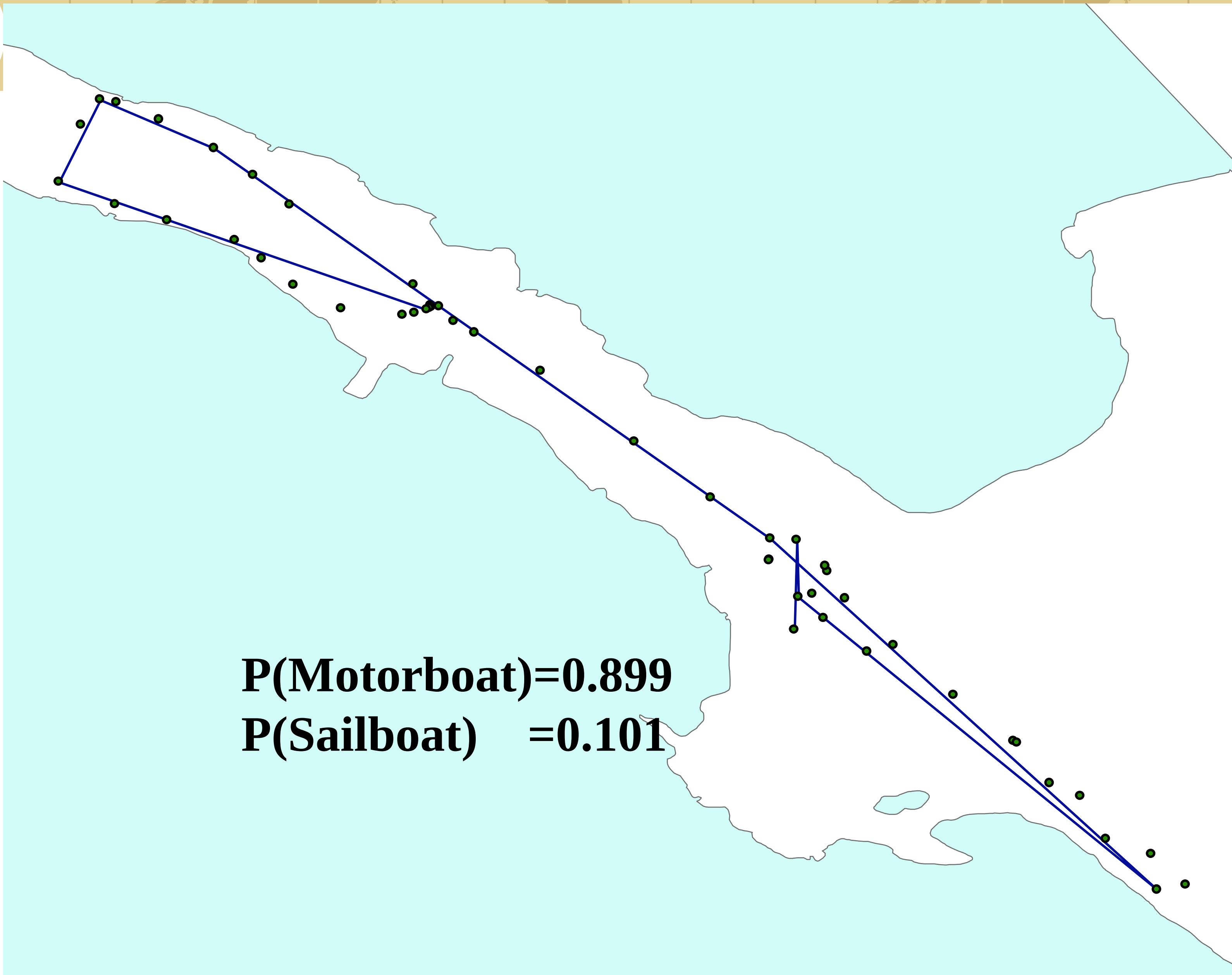
92.6% of original grouped cases correctly classified.

Key results

- **Stepwise discriminant analyses for the coastal and river study areas retain the same independent variables**
  - **Mean Speed, Mean Turning Angle, Farthest distance to shore**
- **Geography does not influence boating trajectories much**
- **Mean Speed, Mean Turning Angle, Farthest distance to shore are significantly different among different groups**



P(Kayak)=0.733
P(Canoe)=0.240


P(Motorboat)=0.899
P(Sailboat)   =0.101

NSERC
CRSNG

Networks of Centres of Excellence
Réseaux de centres d'excellence

THANK YOU!

QUESTIONS?

DALHOUSIE UNIVERSITY
Inspiring Minds

MARIN

CANADA
HYDROGRAPHIC · SERVICE · HYDROGRAPHIQUE